

STIC-Biotech/ChemLib

95965

RECEIVED

JUN -6 2003

(STIC)

From: Ibrahim, Medina A.
Sent: Thursday, June 05, 2003 5:39 PM
T: STIC-Biotech/ChemLib
Subject: 10/090, 035

Please search the following:

1. DNA encoding SEQ ID NO:2 or 4.
2. SEQ ID NO:1-4.
3. oligo of SEQ ID NO:1 and 3

Please search both commercial and patent databases, including pending. Thanks

Medina A. Ibrahim
Patent Examiner
GAU:1638
CM1-9E03
mailbox-9E12
(703)306-5822

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/6/03
Date Completed: 6/15/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: 4
NA Sequences: _____
AA Sequences: 2+2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

reverse to NA
VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 01/05
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library, CM1 - Circ-Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 08:47:15 ; Search time 1860 Seconds
(without alignments)
1455.140 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYIQEVDYCEVRSVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q/cgn2_1/USPTO_spool/US10090035/runat_06062003_105503_10957/app_query.fasta_1.526
-DB-GenEmbl -QFMT=fastap -SUFFIX=rgs -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10090033@cgn_1_1_2566-Runat_06062003_105503_10957 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	295	57.8	173699	8	AP003416 Oryza sat
C 2	295	57.8	174218	2	AC091071 Oryza sat
C 3	295	57.8	185095	8	AP003436 Oryza sat
C 4	295	57.8	185481	2	AC091088 Oryza sat
C 5	225.5	44.2	562	8	SST242802
C 6	93.5	18.3	300000	1	SME591784
C 7	92.5	18.1	814	8	AF003103
C 8	92.5	18.1	815	8	ATAJ2598
C 9	92.5	18.1	1009	8	AY057683
C 10	92.5	18.1	10757	1	AE004561
C 11	92.5	18.1	198780	8	ATCHRIV86
C 12	92.5	18.1	206420	8	ATAP21
C 13	92	18.0	1898	10	MUSDELTA
C 14	92	18.0	2330	10	MUSUCRBP
C 15	92	18.0	3041	10	MUSUCRBP
C 16	91.5	17.9	149879	2	AP005381
C 17	91.5	17.9	163285	2	AP004585
C 18	91	17.8	130814	2	AC121354
C 19	91	17.8	177392	2	AC127187
C 20	90.5	17.7	203050	1	AL646071
C 21	89	17.5	82972	8	NCH24H17
C 22	88.5	17.4	992	10	BC027007
C 23	88.5	17.4	158414	10	AC005960
C 24	88.5	17.4	196050	1	AL646058
C 25	88.5	17.4	197355	2	AC116130
C 26	88	17.3	8874	9	HSREP10
C 27	88	17.3	197050	1	AL646081
C 28	87.5	17.2	163660	2	AC100745
C 29	87	17.1	110000	2	LMFLCHR16_02
C 30	87	17.1	167050	2	AC119146
C 31	87	17.1	171749	2	AC112440
C 32	87	17.1	178665	2	AP003483
C 33	87	17.1	179137	2	AP000853
C 34	87	17.1	184569	9	AC010583
C 35	87	17.1	210872	10	AC084386
C 36	86.5	17.0	1458	10	MUSBDL2A
C 37	86.5	17.0	1766	10	BC016434
C 38	86.5	17.0	4151	10	AF180338
C 39	86.5	17.0	37805	3	AF125964
C 40	86.5	17.0	181419	10	AL591665
C 41	86	16.9	4131	9	AB023233
C 42	86	16.9	63391	8	NC6882
C 43	86	16.9	147481	2	AC067762
C 44	86	16.9	165120	9	AL138704
C 45	86	16.9	192149	9	AC091884

ALIGNMENTS

AP003416/c	173699 bp	DNA	linear	PLN 17-MAY-2002
LOCUS				
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0456E05.			
ACCESSION	AP003416			
VERSION	AP003416.4	GI:20804922		
KEYWORDS	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0456E05.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.			
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0456E05			
JOURNAL	Published Only in Database (2001)			
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-MAR-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan			
COMMENT	(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) On May 15, 2002 this sequence version replaced gi:18181981. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI Nonredundant protein databases with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0456E05 clone has an overlap with P0470A12 (DBJ: AP003436) at the position 130,058 to 173,699 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.			
FEATURES	Location/Qualifiers			
source	1..173699 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clone="P0456E05" join(9483..9641,10723..10785) /gene="P0456E05.1" join(9483..9641,10723..10785) /gene="P0456E05.1" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92602.1" /db_xref="GI:20804923" /translation="MCLMPSTACVCVERQQTNQNLATPPAPAPAAAFALYYVVLASCDQNCVHARVAVWASIIIN" complement(join(11701..11960,11998..12076,12765..12881,13131..13310,13459..13587,13660..13758,14527..14631,15002..15131,15421..15487,17632..17717,17924..17952)) /gene="P0456E05.2" complement(join(11701..11960,11998..12076,12765..12881,13131..13310,13459..13587,13660..13758,14527..14631,15002..15131,15421..15487,17632..17717,17924..17952))			
gene	/gene="P0456E05.2" /note="contains ESTs AU0101913(S0657),AU01912(S0657) similar to rabgap domains protein unknown protein" /codon_start=1 /protein_id="BAB92603.1" /db_xref="GI:20804924" /translation="MHEPLYRVHKKQOPVTAQHRPLRGTEATLGSIRGRRGSPHETGGFNETVWALLKHENFDCSLWVGEALQSLSKVIDLDELMLAAGVDPGAARPTVWKLLGVLPDSRALWEQELAKRKSQYAAKEEFLSNPYSEIIEQIDRDVKRTPDMHFPCDSEFAKSNQESLKNILIEFAKLNAGIRYVQGMNEILAPLFVFRNDPDDKNMFAEADFFCFMELLSGFRDNFCQKLDNSAVGIQGTLSKLISOLVAKYDQELORYLEITFEI NPOFYAFRWITLLLTQEFNADTIHNDILLSDPDQSGRSIYRDGTGSMYNNIETS VAHAYQSSRNSAVPVVEKLHRCRKATGAARQAAPFLSPFRSPPTTENAEAFPC RFLAATRPISVQGDNLNLFPSLPYSCSLAAAGLVL" join(18574..18678,18821..18984,20752..21058,21870..22177,22377..22536) /gene="P0456E05.3" join(18574..18678,18821..18984,20752..21058,21870..22177,22377..22536) /gene="P0456E05.3" /note="contains EST AU070785(R10238) similar to Arabidopsis thaliana chromosome 3, At3g48360 unknown protein" /codon_start=1 /protein_id="BAB92604.1" /db_xref="GI:20804925" /translation="MCEGVRAAGDAAAADVDVITSSGRRIPAHSTVLASASPVLSES ILQRLKTERDAAAGGKVRRAVYLIRGVTDAAAFLVRLIAGSSGDEIDEKSKAA QMLVAHAYRVPWLKRCCEGASIGSLTAESVYDVTMLAALCDAPOLHTRLLAKSF KAVEKTEARFLQENDPWLELDIQLRHADARRRRKRAEQGVYVSEAMDCLS HICVPGTEVPGVRAPAAAPCPAYATACGLQLIRHFSRCHTSCPCRCMOLLIR LHAALCDLPDGHCTPTPLCMQFRKEEKAARAKAKAGDDDDKGLLVKVRVARAMS SLGRRKQMSCQC" complement(join(23558..23638,24837..24941)) /gene="P0456E05.4" complement(join(23558..23638,24837..24941)) /gene="P0456E05.4" /note="contains EST AU162265(E30874) unknown protein" /codon_start=1 /protein_id="BAB92605.1" /db_xref="GI:20804926" /translation="NVKVATYFAMTFGAFLFWQSMRHHVHIALHQDEKKEMEREQE IKRMOALMAQAKES" complement(join(26102..26407,26766..26901,27028..27170,27656..27784,28154..28402,28988..29359)) /gene="P0456E05.5" complement(join(26102..26407,26766..26901,27028..27170,27656..27784,28154..28402,28988..29359)) /gene="P0456E05.5" /note="hypothetical protein magnesium transporter-like protein" /codon_start=1 /protein_id="BAB92606.1" /db_xref="GI:20804927" /translation="MRPSAAGGGGGGRRKAAARAAARENLVVPASGARVEEA GKHAVMARTGLPARDLVLDPLISTEITLGERAIVNLERKAVITPAEVLTPNSK DPAPASFVCDLOARVLASSDQAEEFMGESAVTSFPFALTITPPELEMTNKNRS NVYGGTHSNMPTLTAADGNTKVLPEFFRALEVCESACRSLEETSTLQEPAYPA LDELTSTLNLERVRLQKSLRVAISGRVKVRDEHLDDMDMAEMYLETKLIR QEISTSSRVEVDDPSQLEVDREDDYREADYSNGTFIGYKPHIELEMLLEAFVQI DGTLNKLSHLREYVDDEDITINMLDDKQNLQMGVLSATVTVITAGVAVGLFGM NIGISLYADPTNEEKRASNMKFETTLGTAGCTVWYIVANGVKRSGLLO" join(36977..37082,37255..37479,37607..37701,37823..38015,38424..38764,39169..39399,39493..39645) /gene="P0456E05.6" join(36977..37082,37255..37479,37607..37701,37823..38015,38424..38764,39169..39399,39493..39645) /gene="P0456E05.6" /note="contains ESTs D46801(S11706),AU095929(S11706)" /codon_start=1 /product="proline transport protein-like"			

```

/protein_id="BAB92607.1"
/db_xref="GI:20804928"
/translation="MAASRLDFAAAAGQODKAGSGSYTTAATAHQIDTDSQQQGVGL
LVYGFNCALVFSNMAPDQMGACILFLIGAASWYANMLLAGLHVIGQIRY
RLDMGEYFAFYFPTISAMRWLATSAAVAVDVALIVILIDGSKNSPKQDNVH
GTOAEKVFAGFAIAAILVCNTSGLLPEIQSLTRKPVVNMERALLQYTAGAAVYG
ISVAGWYGAAYSEVLPDQSLSPGWVLIATAFAGISVSQHLFTVPIHEALDTQM
OLDEGMFSRNLGRLLAGLYFGANAFVTFPFMGDFVNLFGSFLPFLTEFPPS
MYLTKTGDEAGRNRIHWGLIIVASSVLSVTTAAAVRLIVHNASYHFADFM
complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/gene="P0456E05.7"
CDS
complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/note="contains ESTs AU101665(R2414),D24709(R2594)
similar to Arabidopsis thaliana chromosome 5, At5g14420
unknown protein"
/codon_start=1
/protein_id="BAB92608.1"
/db_xref="GI:20804929"
/translation="MGGGSRSPRDSGHRGHSYSPFQQQWGGGGGGGGGGGGY
GQDAHGGYGAAPQGGYAAPIYQPPMPSPAATOPPARAGASKPRLDRRYS
RIADHVSVDQVTDALAQAGLESSNLIVIDFTKSNWLGKFSFGSLHHSNAPNP
YEASTIIGTQSLKFDENLIPCFGFGDATTHDQDFCFYDLRPGNFGSEALRYRE
LVPHRLAGPTSFAPLIEMANTIVEQSGGQYHLLIADGQVTRSDVTSQSSQEQ
KTVDIAVASLPLISIVLVGDFGDMWMEKFDNIPSFADNFQFVNFSEIMNKNP
OSRKEAALSALEMIPOQYKATVELGILGRSFSKSPERVLPPPGSHDAYSYGSKS
FSKSTYQSSTSSSPHYETAQSSPAVPSSTYDQVCPICLVNPKMDAFCGGHQC
NPCAFERLADPKPVYWCIDKDKASNPQLDQCGAGDVSSIETH"
/gene="P0456E05.8"
51367..52887
CDS
51367..52887
/gene="P0456E05.8"
/note="hypothetical protein
similar to salt-inducible protein"
/codon_start=1
/protein_id="BAB92609.1"
/db_xref="GI:20804930"
/translation="MAASRLRLKPLSLKHQRLSPEDVQOAEPSASSISIPLD
PSLPILPLVLSLSPSPALPSVHASTPAALRIURRHHPRHAAALDLHLLAA
SDSPAFRDLRLTSLASRLAESRLRLLELVLFPCPTDSDIFACPELLTF
RKALLAFASGDIIPASDALASLRRAADSPFAEFYNIILHALARLRHEDTVRYGE
MTSTHVRHPDVTYFNILINSCEAGVDAAMRWFOEMRWSCPTGVSFNTLMRGFER
EGCKEGIKVAREMLQGLFGLSVASMEIMINGLCRGEGLKAAVEIEELVDGVPEG
FDCLDLVESLCVRNVKAVEVVELILERNVNSCLGVPGVTVLECLMEKGLDKRAC
MGRMAAEIVPDTISCNVIEFALCEGRVDAANRLKLQAKRKGFDAGFTYSIIVOG
FGRQIRKEGEAVLDEMDAGYFVNATYRNRLDGLHMGSRMQLQKCRHNRNAN"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
55150..55176..55310,55441..55668)

```

Alignment Scores:

```

Pred. No.: 3 08e-17 Length: 173699
Score: 295.00 Matches: 57
Percent Similarity: 75.53% Conservative: 14
Best Local Similarity: 60.64% Mismatches: 19
Query Match: 57.84% Indels: 4
DB: 8 Gaps: 2

```

US-10-090-035-2 (1-93) x AP003416 (1-173699)

```

Qy 1 MetaLactTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro---- 19
Db 157571 ATGGCTCACTACCAGAGGTGGACTACTGCTCGGAGAGGTGAGTCTGGTACCCACC 157512
Qy 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluPheGlu 39
Db 157511 GCGCGCTTCTCGCGCGCGCGCGTGCAGCAGCAGCAGTCTGCAAGGAGCGTTCCAG 157452

```

```

Qy 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157451 GAGATCGAC-----AGTCCGGCTCCCGCGCCGCCACCACCAACACCGGC 157401
Qy 60 GlyHisGlyPheValAlaArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157400 AACGACTACCTGATGTTGCGGAGACCAAGGTGGAGGAGACTTCAACACCTGCACCGGC 157341
Qy 80 GluValHisGluArgGluSerPheIleAlaArgAlaAsn 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157340 GAGTCCCGGACGACAGAGTCTCTCTCAAGTCGCAC 157299

```

RESULT 2
LOCUS AC091071 174218 bp DNA linear HTG 27-MAR-2001
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.
ACCESSION AC091071 GI:13450004
VERSION HTG; HTGS PHASE2.
KEYWORDS Oryza sativa (japonica cultivar-group).
SOURCE Oryza sativa (japonica cultivar-group).
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 174218)
AUTHORS Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
and Eun, M.-Y.
TITLE Oryza sativa PAC P0456E05 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174218)
AUTHORS Hahn, J.-H., Eun, M.-Y. and Kim, H.-I.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology (NIAT), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@rda.go.kr,
Tel: 82-31-290-0309, Fax: 82-31-290-0308)
COMMENT The PAC clone was provided by Japanese Rice Genome Research Program
of NIAR.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 53113: contig of 53113 bp in length
* 53114 53213: gap of unknown length
* 53214 172404: contig of 119191 bp in length
* 172405 172504: gap of unknown length
* 172505 174218: contig of 1714 bp in length.

FEATURES

source

```

1..174218
/organism="Oryza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0456E05"

```

BASE COUNT 49037 a 38377 c 37185 g 49417 t 202 others
ORIGIN

Alignment Scores:

```

Pred. No.: 3 08e-17 Length: 174218
Score: 295.00 Matches: 57
Percent Similarity: 75.53% Conservative: 14
Best Local Similarity: 60.64% Mismatches: 19
Query Match: 57.84% Indels: 4
DB: 2 Gaps: 2

```

US-10-090-035-2 (1-93) x AC091071 (1-174218)

/product="putative glycine-rich protein"
/protein_id="CA61838.1"
/db_xref="GI:6478144"
/translation="MAHFKEMDYDEVTSMKPGFGRHGGGVQVQFVVKETEEVEQV
TPRGRSGHHGHRGNOGHGSHFQARETFERHQHARRVPRQAQGERPCQ"

BASE COUNT 143 a 153 c 154 g 112 t
ORIGIN

Alignment Scores:
Pred. No.: 3.68e-13 Length: 562
Score: 225.50 Matches: 59
Percent Similarity: 70.00% Conservative: 11
Best Local Similarity: 59.00% Mismatches: 20
Query Match: 44.22% Indels: 11
DB: 8 Gaps: 6

US-10-090-035-2 (1-93) x SST242802 (1-562)

Qy 1 MetAlaTyrTyrGlnGlu---ValAspTyrCysSerGluGluValArgSerValAlaPro 19
Db 63 ATGGCGCATTCAAGGAGATGGTGACTAC-----GAGGAGGTGACCTCGATGGCCAAA 116
Qy 20 AlaGlyPheGlyArgHis---GlyGlyGlyValGlnGlnHisValValGlyGlyLysPhe 38
Db 117 CCGGCTTCGCGCCGCCGCGCGCGCTCCAGCAGTTCGTCTGCTCAAGGAGACGCTTC 176
Qy 39 GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisHis 58
Db 177 GAGGAGTTCGAGCAGGTCCACCCGCGCGCAGTTCAGGCGCGCAGTTCGAGTTCGAGTTC 233
Qy 59 Gly-----GlyHisGly-----PheValValArgGluThrArgValGluGluAsp 73
Db 234 GCAACCAAGGCGCCGCGCTCCGCGCAGTTCAGGCGCGCGAGACCAAGTTTGAGA--GAC 291
Qy 74 IleAsnThrCysThrGlyGlyValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 292 ATCAACACGCGCAGCGCGAGTTCACGAGGCGAGGAGACGCTCCTGTTCAGGCGCTGAC 351

RESULT 6
SME591784/c 300000 bp DNA linear BCT 05-JUL-2002
LOCUS
DEFINITION Sinorhizobium meliloti 1021 complete chromosome, segment 3/12.
ACCESSION AL591784 AL591688
VERSION AL591784.1 GI:15073438
KEYWORDS
SOURCE Sinorhizobium meliloti.
ORGANISM Sinorhizobium meliloti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group:
Rhizobiaceae; Sinorhizobium.
REFERENCE 1 (bases 1 to 300000)
AUTHORS Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Puehler, A., Purnelle, B., Ramsperger, U., Renard, C., Thebault, P., Vandenbol, M., Weidner, S. and Galibert, F.
TITLE Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)
MEDLINE 21396507
PUBMED 11481430
REFERENCE 2 (bases 1 to 300000)
AUTHORS Gouzy, J.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO EU Consortium
COMMENT
MELILO EU Consortium:
Laboratoire de Biologie Moleculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Genetique et Developpement UMR6061-CNRS,
Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,
D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,

93453 138783: contig of 45331 bp in length
* 138784 138883: gap of unknown length
* 138884 185481: contig of 46598 bp in length.
Location/Qualifiers
1. 185481
/organism="Oryza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0470A12"
BASE COUNT 51439 a 41522 c 41493 g 50653 t 374 others
ORIGIN

Alignment Scores:
Pred. No.: 3.3e-17 Length: 185481
Score: 295.00 Matches: 57
Percent Similarity: 75.53% Conservative: 14
Best Local Similarity: 60.64% Mismatches: 19
Query Match: 57.84% Indels: 4
DB: 2 Gaps: 2

US-10-090-035-2 (1-93) x AC091088 (1-185481)

Qy 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db 156721 ATGGCTCACCACGAGGTGGACTACTCTCGAGGAGGTGAGTGGTGCACCCACC 156780
Qy 20 AlaGlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValGlyGlyLysPheGlu 39
Db 156781 GCGGCTTCCTCGCGCGCGCGCGTGCAGCAGCAGCAGTCTCAGGAGAGTTCACAG 156840
Qy 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db 156841 GAGATCGAC-----AGTCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 156891
Qy 60 GlyHisGlyPheValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 156892 AACGACTACCTGATGGTGGCGGAGACCAAGTGGAGGAGACTTCACACCTCAGCGGC 156951
Qy 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 156952 GAGTCCCGCAGCGAGCAGAGTTCCTGCTCAAGTCCGAC 156993

RESULT 5
SST242802 562 bp mRNA linear PLN 27-NOV-1999
LOCUS
DEFINITION Sinorhizobium meliloti 1021 complete chromosome, segment 3/12.
ACCESSION AJ242802
VERSION AJ242802.1 GI:6478143
KEYWORDS
SOURCE Sinorhizobium meliloti.
ORGANISM Sinorhizobium meliloti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group:
Rhizobiaceae; Sinorhizobium.
REFERENCE 1 (bases 1 to 562)
AUTHORS Neale, A.D., Blomstedt, C.K., Bronson, P., Le, T.N., Guthridge, K., Evans, J., Gaff, D.F. and Hamill, J.D.
TITLE The isolation of lowly-transcribed genes which are induced during desiccation of the resurrection grass Sporobolus stapfianus
JOURNAL Unpublished
MEDLINE 2 (bases 1 to 562)
PUBMED
REFERENCE 2 (bases 1 to 562)
AUTHORS Neale, A.D.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1999) Neale A.D., Biological Sciences, Monash University, Wellington Rd Clayton, Victoria, 3168, AUSTRALIA
COMMENT
Location/Qualifiers
1. 562
/organism="Sporobolus stapfianus"
/db_xref="taxon:56623"
/clone="SDG137c"
63. 350
/codon_start=1

FEATURES
source
CDS

Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte
des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,
B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@toulouse.inra.fr
http://sequence.toulouse.inra.fr/melliloti.html.

FEATURES

```

source
1. .300000
/organism="Sinorhizobium melliloti"
/strain="1021"
/db_xref="taxon:382"
gene
115. .1149
/genes="SMC02146"
CDS
115. .1149
/genes="SMC02146"
/function="cell processes; transport of small molecules"
/notes="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE PHOSPHATE-BINDING PERIPLASMIC PROTEIN"
/protein_id="CAC45081.1"
/db_xref="GI:15073439"
/translation="MKSLLKLVAAALVASAAGFAGVAAADQTOVAGSSVTLVPIAKIVAE
NFEFTPFKPTPVVSEGGSSAGLKEFCCKGVEDIDTANSRKINDSEVEACKAAGVT
ETQAEIKYIGDVGATDAGNADLALKPEHLKALAEIVVDGKLVANPYKKWSEVDAS
LPDAEIAIYIGEKHATREYFEFKILAAAGCKDGAQDVKAASDERQKQSKCVARVK
DGMVADIGDTETILARISNKGIVGFSFYENNADKLKATVNGIVSTETIASG
EYPSRPLFFVYKKAHIGVIPGLKEYVEFFVDQMGIDPSPLAEYGLVAAPDAEREII
RTSFEAGNL"
1241. .1846
/genes="SMC02145"
1241. .1846
/genes="SMC02145"
/function="miscellaneous; hypothetical/partial homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="HYPOTHETICAL SIGNAL PEPTIDE PROTEIN"
/protein_id="CAC45082.1"
/db_xref="GI:15073440"
/db_xref="SPTREMBL:Q92K00"
/translation="MSTSLILLIILIGFTAFIAGRARSPARSGGKLASLHSLPGYHG
SYVAIWMALPALLIALLVWLIAPIIIDYTVRSLPEFPVQNGAATVOLALGVKSVAA
GFERLSADELAELNGSTNNREVLASKVALAGEGEPYMQVQAQFNAMRQYSRWLMS
AVLLVSVAGGLYAIRNIDVFRRAAIGWSASCARSSSRP"
1813. .2724
/genes="pstC OR SMC02144"
1813. .2724
/genes="pstC OR SMC02144"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : putative
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT SYSTEM PERMEASE ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45083.1"
/db_xref="GI:15073441"
/db_xref="SPTREMBL:Q92SA3"
/translation="MUGSLIVASSIAILLTVTGIVASMLSEAGRFTTMSPPFFFTGV
WDPRFAAAGSGAGSQFGLPLLAGTLYIAFVAMLEAFVPIGLFAAIYMAEYASPRLS
TAKPLLEVLACIPITVIVGFPAITVGPFLRDISITQLAGLTVQNTQNFQAGSVITAGI
VAGIMLPFVSSLSDDIITQVPRALRDGSLGATRSETVKYKILPALPGIIVGALLL
TASRAVGETMIVLAAGVAARMNPLEPMPTVTTVTKIVNQLTGDLEFTSQPTLVAFAL
GITLFTVITGLNIYALIVIRKYREQE"
2721. .4043
/genes="pstA OR SMC02143"
2721. .4043
/genes="pstA OR SMC02143"
/function="cell processes; transport of small molecules"
/notes="Product confidence : putative
Gene name confidence : putative
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE PHOSPHATE TRANSPORT SYSTEM PERMEASE ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45084.1"
/db_xref="GI:15073442"
/db_xref="SPTREMBL:Q92SA2"
/translation="MTEFASPTVTVSTRPAAARRDIDGIRRYAAARRFRFAGMSATAI
GLFLVALLMSVSKGYTAFQQTITLPIEFSEKVIDPKNERATDPKOLYTANYPLLV
REALVKGOLDPSKRLPLVKAQADMSASARTQIRDIVADPSVIGRTIPVTLAEANI
DSAFKQIDLSVDETKRKVKDEQVAMNQLAEAGVLKKSFTGIFTFGSSSRPEAGV
GVALGLSVMMLIVLSLPIGVAAISYLEEAPKRLTDLIEVINNNLAAPSVIYG
LIGLAVNFAGPRAALVGFVLTMLPTLIIITRAALKAVPSPISNSAALGLGAS
KMQTIFHVLPLAMPGLITGTIIGLAHAGETAPLLIICWVAFVDFPSPLEPSTAL
PVOIYMWANAEARAFVERISGAILIILLIIFMNLGAILLRFRFRFR"
4057. .4872
/genes="pstB OR SMC02142"
4057. .4872
/genes="pstB OR SMC02142"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : probable
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT ATP-BINDING ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45085.1"
/db_xref="GI:15073443"
/db_xref="SPTREMBL:Q92SA1"
/translation="NNINSEAAVEKALDQKNTPYKMGINDSVSYTGKRALFDVNL
NIRENTVALLIGPSCGKSTFRLNRRNDTIENCRVTGKITLDEDDIYDPSIDVVEL
RARVGNVQKPNFPFYSIDNVSYGPRHLRHAETKAELEDEVETSLQKAGLWNEVKDR
LOEPGTLSGGQQLCLCARAVAVSPEVILMDEFCSDALDPIATARVELIHELRAFET
IVIVTHSQQAARVSQRTAMFHLGNLVEENDTKMFTNPDDQRTQDYIMGRFG"
4959. .5672
/genes="phoU OR SMC02141"
4959. .5672
/genes="phoU OR SMC02141"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : probable
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT SYSTEM
TRANSCRIPTIONAL REGULATOR PROTEIN"
/protein_id="CAC45086.1"
/db_xref="GI:15073444"
/db_xref="SPTREMBL:Q92SA3"
/translation="MSHAHMSAPDEELKYLTRRISDMGLBQMVADSVRALVNSDL

```

```

REFERENCE
AUTHORS      2 (bases 1 to 814)
              Okamuro,J.K., Caster,B., Villarreal,R., Van Montagu,M. and
              Jofuku,K.D.
TITLE        Direct Submission
JOURNAL      Submitted (08-MAY-1997) Biology, University of California, Santa
              Cruz, CA 95064, USA
FEATURES
Source       1. .814
              /organism="Arabidopsis thaliana"
              /db_xref="taxon:3702"
misc_feature 1..814
              /note="similar to EST with GenBank Accession Number
              T76017"
CDS          /note="putative DNA binding protein; similar to A.
              thaliana APTAL2 encoded by Genbank Accession Number
              U12546"
              /codon_start=2
              /product="AP2 domain containing protein RAP2.10"
              /protein_id="AAC49776.1"
              /db_xref="GI:2281645"
              /translation="FFETSSILRFRRYPLLLSSWPTCSSISPSACNFAVNITSSFFFFFIIF

```

[illegible]

Qy	70	VaiglucluaspileasnThrCysThrGlyGlu	80
	:::	:::	:::
Db	665	ATTAATGATGATCTTATGGAGTGTAAGTAAA	697
RESULT	8		
ATAJ2598		815 bp	mRNA
LOCUS	ATAJ2598		linear
DEFINITION	Arabidopsis thaliana mRNA for TINY-like protein.		
ACCESSION	AJ002598		
VERSION	AJ002598.1	GI:2632062	
KEYWORDS	TINY-like protein.		
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 815) Terry,N., Heijnen,L., De Keyser,A., Van Asselidonck,M., De Clercq,R., Verbakel,H., Gielen,J., Zabeau,M., Villarroel,R., Jesse,T., Neyt,P., Hogers,R., Van Den Daele,H., Ardiles,W.,		

COMMENT
FEATURING
this cDNA is corresponding to EST 140D13 (T76017).
interuniversity institute for biotechnology, K.L. Ledegancxstraat
35, 9000 Gent, BELGIUM

Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologos, A. and Ecker, J. R.

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CDNAs (RAFL CDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

Akmal, C.J., Roessena, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
Dale, J.C., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
Leam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, R.C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

```

Location/Qualifiers
1. .1009
   /organism="Arabidopsis thaliana"
   /db_xref="taxon:3702"
   /chromosome="4M"

```

```

/chrosome="4"
/clone="RAFL09-50-B19(R14720)"
/note="ecotype: Columbia"
1..201
202 792

```

```

note: :122
/note="TINY-like protein"
/codon_start=1
/product="AT4g38900/C7A10_460"
/protein_id="BAR1514_1"

```

```

/Protein_ID: mma32j14.1
/db_xref="GI:16323159"
/translation="METATEVATVSPAVTAAVATRKRDKPKYKIGIRMRKWKWVAE
IREPNKRRIWLGSSYPEAAARAYDTAVFYLRSPSARLNFPELLAGVTVTGGGGGV

```

NGGDMSAAYIRRKAAEVGAQVDALEAGAGGNRRHHHHQHQRGNHHDYVDNHSDYRIN
DDLMECSKEGFKRCNGSLERVDLNKLDPETSDDD "

```

/note="compared to genome"
/replace="g"
793..1009
266 a 183 c 269 g 291 t
res:

```

res:

Length.	1000
3.53	

0:00
ending:

[illegible]

exon	12203..12795 /gene="AT4g36750" /number=1	intron	/number=7 15954..16061 /gene="AT4g36760" /number=7
intron	12796..13547 /gene="AT4g36750" /number=1	exon	16062..16128 /gene="AT4g36760" /number=8
exon	13548..13776 /gene="AT4g36750" /number=2	intron	16129..16204 /gene="AT4g36760" /number=8
gene	14283..17541 /gene="AT4g36760" join(14283..14342,14482..14545,14645..14784,14874..15158, 15237..15632,15717..15803,15889..15953,16062..16128, 16205..16339,16440..16520,16618..16672,16768..16872, 17002..17073,17162..17286,17374..17541)	exon	16205..16339 /gene="AT4g36760" /number=9
CDS	15237..15632,15717..15803,15889..15953,16062..16128, 16205..16339,16440..16520,16618..16672,16768..16872, 17002..17073,17162..17286,17374..17541	intron	16340..16439 /gene="AT4g36760" /number=9
	/note="similarity to cytoplasmic aminopeptidase P, Rattus norvegicus, PAX:G2760920 contains EST gb:N96008, AI998300.1" /codon_start=1 /product="aminopeptidase-like protein" /protein_id="CAB80342.1" /db_xref="GI:7270625" /translation="MASHPPLDALVPSSEYHQSEYVSARDKRRFEVSGSGSAGLA LITKKEARLWTDGRIEQLQALQQLSDEWILMRGDEPLVEVWMSDNLPEEANGVDVSWC VSYDTANRKGSKFAKKNQKITTTTDLDEVWKSRRPSPSESPVYVHPLEFAGRSYSHK FEDLRALKEQAGVIEAALDEVAWLYNIRGTDAVCPVWHAFALITDLSAFLYDK KYSDEANSYFNGLVREYTDVSDVALLASDRILISFASKTVOHEAAKMDIEDSD OPRLVNDQMOGLYSGVFLFEAEAKKPKSETSKLTVETVSDKLESLRAKSEHFRG VOYLWLDNOMGLYSGVFLFEAEAKKPKSETSKLTVETVSDKLESLRAKSEHFRG LSPTISSVGSNAVTHYSPEEACAEKDPKTYLDCSGAOYLDGTTDITRVVHFGKP SAHEKECTAVFQGHVAGNAPRPKGTNGYTDILARAPLWKYGLDYRGTHGCVGSY LCVHEGPHQVSRPARNVPLQATMTVDEPGYEDCGNIGRLENLVVNDATETENF GDKGYLQFERITWAPYQVQLDIDELTREIDWLNTYHSKCKDILAPFNQIEMEWLK KATEPVSVSA"	exon	16440..16520 /gene="AT4g36760" /number=10
	14283..14342 /gene="AT4g36760" /number=1	intron	16521..16617 /gene="AT4g36760" /number=10
intron	14343..14481 /gene="AT4g36760" /number=1	exon	16618..16672 /gene="AT4g36760" /number=10
exon	14482..14545 /gene="AT4g36760" /number=2	intron	16673..16767 /gene="AT4g36760" /number=11
intron	14546..14644 /gene="AT4g36760" /number=2	exon	16768..16872 /gene="AT4g36760" /number=11
exon	14645..14784 /gene="AT4g36760" /number=3	intron	16873..17001 /gene="AT4g36760" /number=12
intron	14785..14873 /gene="AT4g36760" /number=3	exon	17002..17073 /gene="AT4g36760" /number=12
exon	14874..15158 /gene="AT4g36760" /number=4	intron	17074..17161 /gene="AT4g36760" /number=13
intron	15159..15236 /gene="AT4g36760" /number=4	exon	17162..17286 /gene="AT4g36760" /number=14
exon	15237..15632 /gene="AT4g36760" /number=5	intron	17287..17373 /gene="AT4g36760" /number=14
intron	15633..15716 /gene="AT4g36760" /number=5	gene	17779..19152 /gene="AT4g36770" complement(17779..19152)
exon	15717..15803 /gene="AT4g36760" /number=6	gene	complement(17779..19152) /gene="AT4g36770" complement(17779..19152)
intron	15804..15888 /gene="AT4g36760" /number=6	CDS	/note="strong similarity to UDP-glucose Glucosyltransferase, Manihot esculenta, PIR2:S41951 Contains UDP-glucosyltransferases signature AA343-386 contains EST gb:AI998300.1" /codon_start=1 /product="glucosyltransferase-like protein" /protein_id="CAB80343.1" /db_xref="GI:7270626" /translation="MELHGVASPGMGHAPVILGLKHLNHHGHDPRVTVFLVTDV SRSKSLIGKTLMEEDPKFVIRFTPLDVSGDLSGLTLKLAEMMRKALPEIKSYMEL EPRPVFVDDLLGTEALEVAKELGIMRKHVLTTSWAFVLAFTVYMASLDKQELYKQLS SIGALLIPGSPVKFERAQDPKRYIRLAEQSQRIGDEVITADGVFNTWHSLEQVTFIG SFDPENLGRVMRGVPVYPVGPLVRPAEPGLKHGVDLDLDLQPKESVYVLLGVGAL TFEQTNELAYGLETHGRFVWVVRPAEDDPSMFDKTKNETEPDLDFPNGLFDRFK DIGLVRTWAPQEEILAHKSTGGFVTHCGNWSVLESIVNGVPMWAMPLYSEOKMNMAM VSGELKATLQINVADGIVKKEVIAEMVKRVMDDEEGKMRKNVKKELKTAEEALNMTH IPSAYFT"
exon	15889..15953 /gene="AT4g36760"		complement(17779..19152)


```

/tissue_lib="129SV Strataqene"

```

```
promoter      1076.  .1808
protein_bind  1728.  .1737
              /bound_moiety="SP1"
exon          1776.  .2885
              /number=1
```

BASE COUNT	676 a	886 c	980 g	499 t
ORIGIN				

Alignment Scores:		
Pred. No.:	12.3	3041
Score:	92.00	
Percent Similarity:	40.85%	
Best Local Similarity:	30.99%	
Query Match:	18.04%	
DA:	10	2
	Length:	3041
	Matches:	22
	Conservative:	7
	Mismatches:	25
	Indels:	16
	Gaps:	2

US-10-090-035-2 (1-93) x MUSTRANS01 (1-3041)

[illegible]

Search completed: June 16, 2003, 11:37:14
Job time : 2119.5 secs

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 496)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.

FEATURES

source

1..496
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: PGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT
ORIGIN

97 a 156 c 153 g 90 t

Query Match

Best Local Similarity 67.5%; Score 456.4; DB 10; Length 496;
Matches 471; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 67 GACACCAAGCGTTCGACCAATGGCTTACTACAGAGGTGGACTTACTGCTCGGAGGA 126
DB 7 GACACCAAGCGTTCGACCAATGGCTTACTACAGAGGTGGACTTACTGCTCGGAGGA 66
QY 127 GGTGAGGTGCGTGGCCCGCGGCTTCGCGCCGACGCGCGGCGGTCCAGCAGCACGT 186
DB 67 GGTGAGGTGCGTGGCCCGCGGCTTCGCGCCGACGCGCGGCGGTCCAGCAGCACGT 126
QY 187 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGGTATCACGCGCGCGGCCAA-CCACACCC 245
DB 127 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGGTCTCACGCGCGCGGCCAATTCATCAC 186
QY 246 ACCACCATGTTCACACGCGCGGCGGCTTCGCGGTGCGGAGACCGGTCGAGGAGG 305
DB 187 ACCACCATGTTCACACGCGCGGCGGCTTCGCGGTGCGGAGACCGGTCGAGGAGG 246
QY 306 ACATCAACACCTCGACGCGGAGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTA 365
DB 247 ACATCAACACCTCGACGCGGAGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTA 306
QY 366 ACTGAGCGCGCGGCGGCGGATCCACGCGGCTTCGCGGTGCGGTGCGGTGCGGTATGTA 425
DB 307 ACTGAGCGCGCGGCGGCGGATCCACGCGGCTTCGCGGTGCGGTGCGGTGCGGTATGTA 366
QY 426 TGTCTGTGTTGACTGTTGTTCAGGGTTCATCGTACTTGGCTATCGTACGTGCACGCACT 485
DB 367 TGTCTGTGTTGACTGTTGTTCAGGGTTCATCGTACTTGGCTATCGTACGTGCACGCACT 426
QY 486 CAGCTCCTGTACGCAATAGCAATAAGCTCGTACCTGAATAAACTTCTTCGTAAT 543
DB 427 CAGCTCCTGTACGCAATAGCAATAAGCTCGTACCTGAATAAACTTCTTCGTAAT 484

Search completed: June 16, 2003, 03:55:49

Job time : 1458.61 secs

The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the p7T3PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

BASE COUNT	ORIGIN
124 a	174 c
159 g	32 t
1 others	

Query Match	68.5%	Score 462.8	DB 12	Length 550
Best Local Similarity	96.8%	Ident. No. 1.4e-41		
Matches	482	Conservative	0	Mismatches 13
			Indels	Gaps
67	QY	GACACACAAAGCGTCTGCACCAATGCTTACTACAGAGAGTGACTACTGCTCGGAGGA	126	
53	Db	GACACACAAAGCGTCGCGACCAATGCTTACTTACAGAGAGTGACTACTGCTCGGAGGA	112	
127	QY	GGTGAGTTCGCTGGCCCGCGCGGCTTCGGCGCCGACAGCGCGCGCGCTCCAGCAGCAGCT	186	
113	Db	GGTGAAGTTCGCTGGCCCGCGCGGCTTCGGCGCCGACAGAGGCGCGCTCCAGCAGCAGCT	172	
187	QY	CGTCAAGGAGAAGTTCGAGAGAGTGCACACGAGTATCACGCGCGCGCGCCA---ACCACCA	243	
173	Db	CGTCAAGGAGAAGTTCGAGAGAGTGCACACGAGTATCACGCGCGCGCGCCA---ACCACCA	232	
244	QY	CCACCACAGTTCACACACGCGCGCACAGGCTTCGTGGTCGCGAGACAGAGGTCGAGGA	303	
233'	Db	CCACCACAGTTCGTCACACGCGCGCACAGGCTTCGTGGTCGCGAGACAGAGGTCGAGGA	292	
304	QY	GGACATCAACACCTGCACCGCGCGAGGTTCACAGAGCGCAGGAGAGCTTCCTCGCCAGGC	363	
293	Db	GGACATCAACACCTGCACCGCGCGAGGTTCACAGAGCGCAGGAGAGCTTCCTCGCCAGGC	352	
364	QY	TAACTGACGCGCGCGCGCGCGGATCCACGCGCGGTTTCGTGGTCGCTGCGTTCGCTTATG	423	
353	Db	TAACTGAGCGCGCGCGCGCGGATCCACGCGCGGTTTCGTGGTCGCTGCGTTCGCTTATG	412	
424	QY	TATGTCGTGGTTGACTGGTTGTCAGGGTCATCGTACTTGGGTATCGTACGTGACCGCA	483	
413	Db	TATGTCGTGGTTGACTGGTTGTCAGGGTCATCGTACTTGGGTATCGTACGTGACCGCA	472	
484	QY	CTCAGCTCCTGTACGAATATACGACAATAAGCTCGTACCTGGAATAAAACTTCTCGTAAAT	543	
473	Db	CTCAGCTCCTGTACGAATATACGACAATAAGCTCGTACCTGGAATAAAACTTCTCGTAAAT	532	
544	QY	ACTAATAACATACATCAAA	561	
533	Db	CGCTTAAAAAATAAAAAA	550	

```

RESULT 14
  A1855425/c
  LOCUS
  DEFINITION
    A1855425 539 bp mRNA linear EST 15-JUL-1999
    603016F02.xl 603 - stressed root cDNA library from Wang/Bohnert lab
    Zea mays cDNA, mRNA sequence.
  ACCESSION
    A1855425
  VERSION
    A1855425.1 GI:5499558
  KEYWORDS
    EST.
  SOURCE
    Zea mays.
  ORGANISM
    Zea mays
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
    clade; Panicoideae; Andropogoneae; Zea.
    1 (bases 1 to 539)
  REFERENCE
    Walbot,V.
    Maize ESTs from various cDNA libraries sequenced at Stanford
  AUTHORS
    University
  TITLE
    Unpublished (1999)
  JOURNAL
    Contact: Walbot V
  COMMENT

```

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.

Location/Qualifiers
1. .539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603." stressed root cDNA library from
Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector:
pBluescriptIII SK(+) XR;
Seeding stressed root cDNA library from Wang/Bohnert lab"
157 c 172 q 119 t
91 a

Query Match	68.0%;	Score 460;	DB 9;	Length 539;
Best Local Similarity	97.6%;	pred. No. 2.9e-41;		
Matches 488;	Conservative 0;	Mismatches 8;	Indels 4;	Gaps 2;
QY	67	GACACACAAGCGTCTGCACCAATGGCTTACTACCAAGGAGGTGGACTACTGCTCGGAGGA	126	
DB	499	GACACACAAGCGTCGGCCACCAATGGCTTACTACCAAGGAGGTGGACTACTGCTCGGAGGA	440	
QY	127	GGTGAGGTCTGTTGGCCCGCGCGCTTTCGGCGCGCCACGGCGCGCGGTCCACGACGACGT	186	
DB	439	GGTGAAGTCTGGTTCGGCCCGCGCGCTTTCGGCGCGCCACGGAGGCGGCTCCAGCAGCAGT	380	
QY	187	CGTCAAGGAGAGTTTCGAGGAGTTCGACAGCGTATCACGCGCGCGCGCCA--ACACCA	243	
DB	379	CGTCAAGGAGAGTTTCGAGGAGTTCGACAGCGTATCACGCGCGCGCGCCAACACCA	320	
QY	244	CCACACCAATGTTACACACGGCGGCCACGGCTTCGTGTCGGAGACACAGGTCGAGGA	303	
DB	319	CCACCACCAATGTTACACACGGCGGCCACGGCTTCGTGTCGGAGACACAGGTCGAGGA	260	
QY	304	GGACATCAACACTCTCACCGCGGAGGTCCACGAGCGCAGGAGAGCTTCTCGCCACGGCG	363	
DB	259	GGACATCAACACTCTCACCGCGGAGGTCCACGAGCGCAGGAGAGCTTCTCGCCACGGCG	200	
QY	364	TAACTGAGCGCCCGCGCGCGCGCATCACGCCGTCGTCTGCTCGCTCGCTGCTTCGTATG	423	
DB	199	TAACTGAGCGCCCGCGCGCGCGCATCACGCCGTCGTCTGCTCGCTCGCTGCTTCGTATG	140	
QY	424	TATGTCGTGGTTGACTGGTGTGTCGAGGTCATCGTACTTGGCTATCGTACGTGCACGCA	483	
DB	139	TATGTCGTGGTTGACTGGTGTGTCGAGGTCATCGTACTTGGCTATCGTACGTGCACGCA	80	
QY	484	CTCAGTCTCTGTACGAATTTACGACAATAAGCTCTGTGACCTGAATTAACCTTCTTCGTAAT	543	
DB	79	CTCACCTCC--GTACGAATTTACGACAATAAGCTCTGTGACCTGAATTAACCTTCTTCGTAAT	21	
QY	544	ACTAATAACCTACATCAAAAA	563	
DB	20	ACTAATAACCTAAAAA	1	

RESULT 15	AW787314	496 bp	mrna	linear	EST
LOCUS	94500206.X1 945 -	Mixed adult tissues from Walbot lab, same as 707			
DEFINITION	(SK) Zea mays CDNA, mRNA sequence.				
ACCESSION	AW787314				
VERSION	AW787314.1				
KEYWORDS	GI:7844111				
SOURCE	EST.				
ORGANISM	Zea mays.				
	Zea mays.				
	Zea mays				


```
REFERENCE 1 (bases 1 to 553)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,
TITLE Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
JOURNAL Maize Mapping Project/DuPont Consensus Sequences for Design of
REFERENCE Overgo Probes
AUTHORS 2 (bases 1 to 553)
TITLE Coe,E.C.
JOURNAL Direct Submission
REFERENCE 25-APR-2002 Maize Mapping Project, University of
AUTHORS Submitted (Columbia, MO 65211, USA
JOURNAL Missouri, Columbia, MO 65211, USA
FEATURES Location/Qualifiers
source
1..553
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 128 a 173 c 162 g 90 t
ORIGIN
Query Match 70.18; Score 474.2; DB 11; Length 553;
Best Local Similarity 94.38; Pred. No. 8.5e-43;
Matches 516; Conservative 0; Mismatches 23; Indels 8; Gaps 2;
QY 25 GCAGCAATCCACCAAGCACTTCGAAGGACCACTG-----CTCGGAGGACACCAAGCG 79
DB 7 GCAGCAATCCACCAAGCACTTCGACGTCGCACGCGCGCTGCACACAGACACCAAGCG 66
QY 80 TCTGCACCAATGGCTTACTACAGAGGTGACTACTGCTCGGAGGAGGTGAGTCTGCTG 139
DB 67 TCGGCACCAATGGCTTACTACAGAGGTGACTACTGCTCGGAGGAGGTGAGTCTGCTG 126
QY 140 GCCCGCGCGGCTTCGGCGCGCACGCGCGCGCTCCAGCAGCAGCTGCTCAAGAGGAAG 199
DB 127 GCCCGCGCGGCTTCGGCGCGCACGCGCGCGCTCCAGCAGCAGCTGCTCAAGAGGAAG 186
QY 200 TTCGAGGAGGTGCACACGCTATCAGCGCGCGCGCGCA---ACGACCAACCAACCATGGT 256
DB 187 TTCGAGGAGGTGCACACGCTATCAGCGCGCGCGCGCAACCAACCAACCAACCATGGT 246
QY 257 CACCAAGCGCGCGCGCTTCGCTGTCGCGAGACAGGCTCGAGGAGGACATCAACACC 316
DB 247 CACCAAGCGCGCGCGCTTCGCTGTCGCGAGACAGGCTCGAGGAGGACATCAACACC 306
QY 317 TCACCGCGGAGGTCCACAGCGCAGGAGAGCTTCCTCGCGAGGGCTAACTGAGCGGCC 376
DB 307 TCACCGCGGAGGTCCACAGCGCAGGAGAGCTTCCTCGCGAGGGCTAACTGAGCGGCC 366
QY 377 CGCGCGCGCGATCCACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
DB 367 CGCGCGCGCGATCCACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
QY 437 GACTGCTGTCAGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB 427 GACTGCTGTCAGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
QY 497 CGAATTAACACATAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
DB 487 CGAATTAACACATAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
QY 557 TCAAAA 563
DB 547 AAAAAA 553
```

```
RESULT 11
AW787732 500 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787732
VERSION AW787732.1 GI:7844510
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 500)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
FEATURES Location/Qualifiers
source
1..500
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 97 a 159 c 156 g 88 t
ORIGIN
Query Match 69.68; Score 470.6; DB 10; Length 500;
Best Local Similarity 99.28; Pred. No. 2.2e-42;
Matches 473; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 67 GACACACCAAGCTTCGACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 126
DB 13 GACACACCAAGCTTCGACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 72
QY 127 GGTGAGGTTCGCTGCGCGCGCTTCGGCGCGCGCGCGCGCTCCAGCAGCAGCT 186
DB 73 GGTGAGGTTCGCTGCGCGCGCTTCGGCGCGCGCGCGCGCTCCAGCAGCAGCT 132
QY 187 CGTCAAGGAGAGTTCGAGAGGTTCGACACGGTATCAGCGCGCGCGCGCGCGCGCGCG 246
DB 133 CGTCAAGGAGAGTTCGAGAGGTTCGACACGGTATCAGCGCGCGCGCGCGCGCGCGCG 192
QY 247 CCACCATGCTCACCACCGCGCGCGCTTCGTGTCGCGCGCGCGCGCGCGCGCGCGCG 306
DB 193 CCACCATGCTCACCACCGCGCGCGCTTCGTGTCGCGCGCGCGCGCGCGCGCGCGCG 252
QY 307 CATCAACACCTGCACCGCGCGAGTCCACAGCGCAGGAGGAGCTTCCTCGCCAGGGCTAA 366
DB 253 CATCAACACCTGCACCGCGCGAGTCCACAGCGCAGGAGGAGCTTCCTCGCCAGGGCTAA 312
QY 367 CTGAGCGCGCGCGCGCGCGCTTCACCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 313 CTGAGCGCGCGCGCGCGCGCTTCACCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCT 372
```


/note="Organ: tassal, kernal, silk, husk, root, leaf;
 Vector: pGAD10; Site.1: EcoRI; cDNA library from fully
 differentiated maize tissues from an active Mutator plant.
 Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
 root, leaf). Unidirectionally cloned. New library number
 given to library 707 for additional sequencing."

```

BASE COUNT      135 a   165 c   160 g   92 t
ORIGIN
Query Match      72.9%; Score 492.6; DB 10; Length 552;
Best Local Similarity 97.3%; Pred. No. 8.9e-45;
Matches 501; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 67 GACACCAAGCGCTTCGACCAATGGCTTACTACGAGGAGTGGACTACTGCTGGAGGA 126
Db 26 GACACCAAGCGCTTCGACCAATGGCTTACTACGAGGAGTGGACTACTGCTGGAGGA 85
QY 127 GGTGAGGTCTGGTGGCGCGCGGCTTCGGCGCGCCACGCGCGGCGGCTCCAGCAGCAGT 186
Db 86 GGTGAGGTCTGGTGGCGCGCGGCTTCGGCGCGCCACGCGCGGCGGCTCCAGCAGCAGT 145
QY 187 GGTCAAGGAGAGTTTCGAGGAGTTCGACACCGGTATACGCGCGCGCGCCACCA 246
Db 146 GGTCAAGGAGAGTTTCGAGGAGTTCGACACCGGTATACGCGCGCGCGCCACCA 205
QY 247 CCACCATGGTTCACACGCGCGCGGCTTCGTTGGTGGCGGAGACGAGGTCGAGGAGGA 306
Db 206 CCACCATGGTTCACACGCGCGCGGCTTCGTTGGTGGCGGAGACGAGGTCGAGGAGGA 265
QY 307 CATCAACACCTTCACCGCGCGGAGTCCACGAGCGGAGGAGGCTTCCTCGCGCGGCTAA 366
Db 266 CATCAACACCTTCACCGCGCGGAGTTCACGAGCGGAGGAGGCTTCCTCGCGCGGCTAA 325
QY 367 CTGAGCGCGCGCGCGGCGGATCCACGCGCGGTTCTGCTGGCTGGCTGCTTATGAT 426
Db 326 CTGAGCGCGCGCGCGGCGGATCCACGCGCGGTTCTGCTGGCTGGCTGCTTATGAT 385
QY 427 GTCTGTGTTGACTGTTGTTCAGGTCATCGTACTTGGCTATGCTGCTGCTGCTATGAT 486
Db 386 GTCTGTGTTGACTGTTGTTCAGGTCATCGTACTTGGCTATGCTGCTGCTGCTATGAT 445
QY 487 AGTCTCTGTACGAATTACGACAATAAGCTCGTACCTGAATAAACTTCTTCGTAATCT 546
Db 446 AGTCTCTGTACGAATTACGACAATAAGCTCGTACCTGAATAAACTTCTTCGTAATCT 505
QY 547 AATACCTACATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 581
Db 506 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540

```

```

RESULT 9
AW288876      524 bp   mRNA   linear   EST 16-JAN-2000
LOCUS      707009E07.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea
DEFINITION   mays cDNA, mRNA sequence.
ACCESSION   AW288876
VERSION      AW288876.1 GI:6695663
KEYWORDS     EST.
SOURCE       Zea mays.
ORGANISM     Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227

```

```

Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.
Location/Qualifiers
1..524
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
)
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/lab_stage="adult"
/lab_host="DH10B"
/note="Organ: tassal, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site.1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT      107 a   170 c   157 g   89 t   1 others
ORIGIN

```

```

Query Match      71.2%; Score 481.2; DB 10; Length 524;
Best Local Similarity 96.5%; Pred. No. 1.5e-43;
Matches 503; Conservative 0; Mismatches 14; Indels 4; Gaps 1;
QY 24 CGCAGCAATCCACACAGCAGCTTCGAGGAGGACCACTGCTCGGAG----GACACACCAAGCG 79
Db 4 CACAGCAATCCACACAGCAGCTTCGAGGAGGACCACTGCTCGGAGGAGGAGGAGGAGGAGG 63
QY 80 TCTGCACCAATGCTTACTACCAAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGAGTGG 139
Db 64 TCGGACCAATGCTTACTACCAAGGAGTGGACTACTGCTCGGAGGAGGAGTGGAGTGGAGTGG 123
QY 140 GCGCGCGCGGCTTCGCGCGCCACGCGCGGCGGCTCCAGCAGCAGCTCTCAAGGAGAGAG 199
Db 124 GCGCGCGCGGCTTCGCGCGCCACGCGCGGCGGCTCCAGCAGCAGCTCTCAAGGAGAGAG 183
QY 200 TTCGAGGAGTTCGACACCGGTATCACGCGCGCGGCGGCGGCTCCAGCAGCAGCTCTCAAGGAGAG 259
Db 184 TTCGAGGAGTTCGACACCGGTATCACGCGCGCGGCGGCGGCTCCAGCAGCAGCTCTCAAGGAGAG 243
QY 260 CACGCGCGCGGCTTCGCGCGCCACGCGCGGCGGCTCCAGGAGGAGGAGGAGGAGGAGGAG 319
Db 244 CACGCGCGCGGCTTCGCGCGCCACGCGCGGCGGCTCCAGGAGGAGGAGGAGGAGGAGGAG 303
QY 320 ACCGCGGAGTTCACGAGCGCAGGAGGAGCTTCCTCGCGCGGCTTAACCTAGCAGCGCGCGG 379
Db 304 ACCGCGGAGTTCACGAGCGCAGGAGGAGCTTCCTCGCGCGGCTTAACCTAGCAGCGCGCGG 363
QY 380 CGGCGGCGATCCACGCGCGGCTTCGCTTGCCTTCGCTGCTGCTTATGATGCTGCTGCTGCTGAC 439
Db 364 CGGCGGCGATCCACGCGCGGCTTCGCTTGCCTTCGCTGCTGCTTATGATGCTGCTGCTGAC 423
QY 440 TGGTGTGCGAGGTCATCGTACTTGGCTATCGTACGTGCGAGCAGCAGCTAGCTCCTGTACGA 499
Db 424 TGGTGTGCGAGGTCATCGTACTTGGCTATCGTACGTGCGAGCAGCAGCTAGCTCCTGTACGA 483
QY 500 ATTACGACAATAAGCTCGTACCTGAATAAACTTCTTCGT 540
Db 484 ATTACGACAATAAGCTCGTACCTGAATAAACTTCTTCGT 524

```

```

RESULT 10
AW104409      553 bp   mRNA   linear   HTC 25-MAY-2002
LOCUS      PC0124784 mRNA sequence.
DEFINITION   Zea mays
ACCESSION   AW104409
VERSION      AW104409.1 GI:21207487
KEYWORDS     HTC.
SOURCE       Zea mays.
ORGANISM     Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

```

```

FEATURES
source
1. 552
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"

```

```

Db      540 ACTAAAAA... 599
Qy      604 AAAAAA 610
Db      600 AAAAAA 606

RESULT 5
BQ619383
LOCUS
DEFINITION
  RNOSEQ7B02_SK.ab1 Salt stressed Zea mays roots cDNA library EST 27-JUN-2002
  mays cDNA clone RNOSEQ7B02_SK.ab1 similar to No homology, mRNA
  sequence.
ACCESSION
  BQ619383
VERSION
  BQ619383.1 GI:21621377
KEYWORDS
  EST.
SOURCE
  Zea mays.
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
  clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
  Wang, H. and Bohnert, H.J.
  Genomics of plant stress tolerance
  Unpublished (2002)
  Contact: Mark Fredrickson
  Department of Plant Biology
  University of Illinois
  1201 W. Gregory Dr., Urbana, IL 61801, USA
  Tel: 2172655473
  Email: bohnertlab@life.uiuc.edu.
FEATURES
  source
  1. 648
  /organism="Zea mays"
  /db_xref="taxon:4577"
  /clone="RNOSEQ7B02_SK.ab1"
  /clone_lib="Salt stressed Zea mays roots cDNA library"
  /tissue_type="Roots"
  /dev_stage="2 weeks old"
  /note="Vector: pBluescript SK+; Stressed 24 hours at 150
  mM NaCl"
  172 a 193 c 183 g 100 t

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN
  Query Match 75.2%; Score 508.6; DB 14; Length 648;
  Best Local Similarity 96.9%; Pred. No. 1.5e-46;
  Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

Qy      67 GACACCAAGCGTCTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 126
Db      60 GACACCAAGCGTGGCCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 119

Qy      127 GGTGAGTGTGGTGGCCCGCGGCTTCGGCCGCGACGGCGGCGGTCCAGCAGCAGT 186
Db      120 GGTGAGTGTGGTGGCCCGCGGCTTCGGCCGCGACGGGTATACAGCGCGCGGCCCA 179

Qy      187 CGTCAAGGAGAGTTCGAGGAGGTTCGACAGCGGTATCACGCGCGCGGCCCA---ACCACCA 243
Db      180 CGTCAAGGAGAGTTCGAGGAGGTTCGACAGCGGTATCACGCGCGCGGCCCA---ACCACCA 239

Qy      244 CCACCAACACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 303
Db      240 CCACCAACACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 299

Qy      304 GGACATCAACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 363
Db      300 GGACATCAACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 359

Qy      364 TAACTGAGCGCGCGCGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 423
Db      360 TAACTGAGCGCGCGCGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 419

Qy      424 TATGCTGTGGTTGACTGTTGTGCGAGGTCATCTACTTGGCTATCTGTCAGTGTCAGCA 483

```

```

Db      420 TATGCTGTGGTTGACTGTTGTGTCAGGGTCATCTACTTGGCTATCTGTCAGTGTCAGCA 479
Qy      484 CTCAGCTCCTGTACGAATTAAGCTGTGACCTGAATGAATAAACTTCTTCGTAAT 543
Db      480 CTCAGCTCCTGTACGAATTAAGCTGTGACCTGAATGAATAAACTTCTTCGTAAT 539

Qy      544 ACTAATACCTACATCAATCAAAAAA... 603
Db      540 ACTAATAAAAAA... 599

Qy      604 AAAAAA 610
Db      600 AAAAAA 606

RESULT 6
BQ619390
LOCUS
DEFINITION
  RNOSEQ7B09_SK.ab1 Salt stressed Zea mays roots cDNA library EST 27-JUN-2002
  mays cDNA clone RNOSEQ7B09_SK.ab1 similar to No homology, mRNA
  sequence.
ACCESSION
  BQ619390
VERSION
  BQ619390.1 GI:21621384
KEYWORDS
  EST.
SOURCE
  Zea mays.
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
  clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
  Wang, H. and Bohnert, H.J.
  Genomics of plant stress tolerance
  Unpublished (2002)
  Contact: Mark Fredrickson
  Department of Plant Biology
  University of Illinois
  1201 W. Gregory Dr., Urbana, IL 61801, USA
  Tel: 2172655473
  Email: bohnertlab@life.uiuc.edu.
FEATURES
  source
  1. 648
  /organism="Zea mays"
  /db_xref="taxon:4577"
  /clone="RNOSEQ7B09_SK.ab1"
  /clone_lib="Salt stressed Zea mays roots cDNA library"
  /tissue_type="Roots"
  /dev_stage="2 weeks old"
  /note="Vector: pBluescript SK+; Stressed 24 hours at 150
  mM NaCl"
  172 a 193 c 183 g 100 t

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN
  Query Match 75.2%; Score 508.6; DB 14; Length 648;
  Best Local Similarity 96.9%; Pred. No. 1.5e-46;
  Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

Qy      67 GACACCAAGCGTCTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 126
Db      60 GACACCAAGCGTGGCCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 119

Qy      127 GGTGAGTGTGGTGGCCCGCGGCTTCGGCCGCGACGGCGGCGGTCCAGCAGCAGT 186
Db      120 GGTGAGTGTGGTGGCCCGCGGCTTCGGCCGCGACGGGTATACAGCGCGCGGCCCA 179

Qy      187 CGTCAAGGAGAGTTCGAGGAGGTTCGACAGCGGTATCACGCGCGCGGCCCA---ACCACCA 243
Db      180 CGTCAAGGAGAGTTCGAGGAGGTTCGACAGCGGTATCACGCGCGCGGCCCA---ACCACCA 239

Qy      244 CCACCAACACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 303
Db      240 CCACCAACACCTGACCGCGCGGCGGTTCGCGTTCGCGGAGACAGGAGGTTCGAGGA 299

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6D01.SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 67 GACACCAACGCTGTCACCAATGGCTTACTACAGAGGTGGACTACTCTCGGAGGA 126
DB 60 GACACCAACGCTGCGCACCACCAATGGCTTACTACAGAGGTGGACTACTCTCGGAGGA 119

QY 127 GGTGAGTCTGTCGCGCGCGGCTTCGGCGCCACGCGCGCGCTCCAGCAGCAGCT 186
DB 120 GGTGAGTCTGTCGCGCGCGGCTTCGGCGCCACGCGCGCTCCAGCAGCAGCT 179

QY 187 CGTCAAGGAGAGTTTCGAGAGGTTCGACACGGTATACGCGCGCGCGCTCCAGCAGCAGCT 243
DB 180 CGTCAAGGAGAGTTTCGAGAGGTTCGACACGGTATACGCGCGCGCGCTCCAGCAGCAGCT 239

QY 244 CCACACCATGTCACCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 303
DB 240 CCACACCATGTCACCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 299

QY 304 GGACATCAACCTGTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 363
DB 300 GGACATCAACCTGTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 359

QY 364 TAATCTGAGCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCT 423
DB 360 TAATCTGAGCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCT 419

QY 424 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 483
DB 420 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 479

QY 484 CTCAGCTCTGTACGAATTTACGACAAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 480 CTCAGCTCTGTACGAATTTACGACAAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539

QY 544 ACTAATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 603

RESULT 4
BQ619337

LOCUS BQ619337 648 bp mRNA linear EST 27-JUN-2002
DEFINITION RNOSEQ6E12.SK.ab1 Salt-stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSEQ6E12.SK.ab1 similar to NO homology, mRNA sequence.

ACCESSION BQ619337
VERSION BQ619337.1
KEYWORDS GI:21621331
SOURCE EST.
ORGANISM Zea mays.

REFERENCE 1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6D01.SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 67 GACACCAACGCTGTCACCAATGGCTTACTACAGAGGTGGACTACTCTCGGAGGA 126
DB 60 GACACCAACGCTGCGCACCACCAATGGCTTACTACAGAGGTGGACTACTCTCGGAGGA 119

QY 127 GGTGAGTCTGTCGCGCGCGGCTTCGGCGCCACGCGCGCGCTCCAGCAGCAGCT 186
DB 120 GGTGAGTCTGTCGCGCGCGGCTTCGGCGCCACGCGCGCTCCAGCAGCAGCT 179

QY 187 CGTCAAGGAGAGTTTCGAGAGGTTCGACACGGTATACGCGCGCGCGCTCCAGCAGCAGCT 243
DB 180 CGTCAAGGAGAGTTTCGAGAGGTTCGACACGGTATACGCGCGCGCGCTCCAGCAGCAGCT 239

QY 244 CCACACCATGTCACCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 303
DB 240 CCACACCATGTCACCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 299

QY 304 GGACATCAACCTGTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 363
DB 300 GGACATCAACCTGTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 359

QY 364 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCT 423
DB 360 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCT 419

QY 424 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 483
DB 420 TAATCTGAGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCT 479

QY 484 CTCAGCTCTGTACGAATTTACGACAAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 480 CTCAGCTCTGTACGAATTTACGACAAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539

QY 544 ACTAATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 603
DB 540 ACTAATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 599

QY 604 AAAAAA 610
DB 600 AAAAAA 606

BASE COUNT	172 a	193 c	183 g	100 t
/clone_lib="Salt stressed Zea mays roots cDNA library"				
/tissue_type="Roots"				
/dev_stage="2 weeks old"				
/note="Vector: pBluescript SK+; Stressed 24 hours at 150				
mm NaCl"				

Query Match	75.2%	Score 508.6	DB 14	Length 648
Best Local Similarity	96.9%	Prod. No. 1.5e-46		
Matches 530	Conservative 0	Mismatches 14	Indels 3	Gaps 1
Qy 67	GACACACCAAGCGTCTCAGCAATGGCTTACTACACAGAGGTGGACTACTGCTCGGAGGA	126		
Db 60	GACACACCAAGCGTCGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA	119		
Qy 127	GGTGAGGTCCGTGGCCCCCGCGGCTTTCGGCCGCCCACGGCGGCGGCGTCCAGCAGCACGT	186		
Db 120	GGTGAGGTCCGTGGCCCCCGCGGCTTTCGGCCGCCCACGGAGGGGCGCTCCAGCAGCACGT	179		
Qy 187	CGTCAAGGAGAAGTTCGAGGAGGTTCGACACGGGTATCACGCGCGGCGGCCA--ACCA	243		
Db 180	CGTCAAGGAGAAGTTCGAGGAGGTTCGACACGGTCTCACGCGCGGCGGCCAACCA	239		
Qy 244	CCACCACCATGTTCACACGGCGGCCACGGCTTCGTGGTGGCGGAGACAGGTCGAGGA	303		
Db 240	CCACCACCATGTTCACACGGCGGCCACGGCTTCGTGGTGGCGGAGACAGGTCGAGGA	299		
Qy 304	GGACATCAACACTGCACCGGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	363		
Db 300	GGACATCAACACTGCACCGGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	359		
Qy 364	TAACTCAGCGCCCGCGCGCCGATCCACGCGCCGTCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	423		
Db 360	TAACTCAGCGCCCGCGCGCCGATCCACGCGCCGTCACGCGCCGTCACGAGGTTCCTCGCCAGGGC	419		
Qy 424	TATGTCGTGGTTGACTGGTTGTCAGGGTCACTCGTACTTGGCTATCGTACGTGCACGCA	483		
Db 420	TATGTCGTGGTTGACTGGTTGTCAGGGTCACTCGTACTTGGCTATCGTACGTGCACGCA	479		
Qy 484	CTCAGCTCTGTACGAATTTACGACATAAGCTCGTCACCTGANTAAACTTCTTCGTAAT	543		
Db 480	CTCAGCTCTGTACGAATTTACGACATAAGCTCGTCACCTGANTAAACTTCTTCGTAAT	539		
Qy 544	ACTAATACCTACATCAAAAAAANAANAANAANAANAANAANAANAANAANAANAANA	603		
Db 540	ACTAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA	599		
Qy 604	AAAAAA 610			
Db 600	AAAAAA 606			

RESULT 2						
B0619315						
LOCUS	BQ619315	648 bp	mrna	linear	EST 27-JUN-2002	
DEFINITION	RNOSEQC06_SK.ab1 salt stressed Zea mays roots cDNA library Zea mays cDNA, clone RNOSEQC06_SK.ab1 similar to No homology, mrna					

VERSION BQ619315.1 GI:21621309
KEYWORDS EST.

SOURCE	ORGANISM
zea mays.	
zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC. clade; Panicoideae; Andropogoneae; Zea.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 648)	Wang, H. and Bohner, H. J.	Genomics of plant stress tolerance	unpublished (2002)
		Contact: Mark Fredricksen	Department of Plant Biology	

```

University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

FEATURES
    source
        1..648
            /organism="Zea mays"
            /db_xref="taxon:4577"
            /clone_lib="RNOSEQ06C06_SK_ab1"
            /clone_lib="Salt stressed Zea mays roots cDNA library"
            /tissue_type="Roots"
            /dev_stage="2 weeks old"
            /note="Vector: pBluescript SK+; Stressed 24 hours at 150
            mM NaCl"
BASE COUNT
    172 a    193 c    183 g    100 t

```

Query Match	75.28;	Score 508.6;	DB 14;	Length 648;
Best Local Similarity	96.9%;	Prod. No. 1.5e-46;		
Matches 530;	Conservative 0;	Mismatches 14;	Indels 3;	Gaps 1;
QY	67	GACACACCAAGCGTCTGCACCAATGGCTTACTACCAAGAGGTGGACTACTGCTCGGAGGA	126	
Db	60	GACACACCAAGCGTGGCACCACCAATGGCTTACTACCAAGAGGTGGACTACTGCTCGGAGGA	119	
QY	127	GGTGAGGTGGTGGCCCCCGCGGCTTCGGCCCGCCACGGCGGGCGGCGCTCCAGCAGCAGCT	186	
Db	120	GGTGAGGTGGTGGCCCCCGCGGCTTCGGCCCGCCACGGAGGGCGGCTCCAGCAGCAGCT	179	
QY	187	CGTCAAGGAGAAGTTTCGAGGAGGTTCGACACGGGTATCACGCCGCGGCGCCA---ACCACCA	243	
Db	180	CGTCAAGGAGAAGTTTCGAGGAGGTTCGACACGGGTCTCACGCCGCGGCGCCAACCAACCA	239	
QY	244	CCACCACCATGTTACCAACGGCGGCCACGGCTTCGTGGTCCGGCGAGACAGGTCGAGGA	303	
Db	240	CCACCACCATGTTACCAACGGCGGCCACGGCTTCGTGGTCCGGCGAGACAGGTCGAGGA	299	
QY	304	GGACATCAACACTTCACCGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	363	
Db	300	GGACATCAACACTTCACCGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	359	
QY	364	TAACTGAGCGCCCGCGCGCGCATCCAGCCCGTTCGTGGTGGCTGGCTGGTGGCTGGCTTATG	423	
Db	360	TAACTGAGCGCCCGCGCGCGCATCCAGCCCGTTCGTGGTGGCTGGCTGGCTGGCTTATG	419	
QY	424	TATGTCGTGGTGTACTGGTGTGCGAGGTTCATCGTACTTGGCTATCGTACGTGCACGCA	483	
Db	420	TATGTCGTGGTGTACTGGTGTGCGAGGTTCATCGTACTTGGCTATCGTACGTGCACGCA	479	
QY	484	CTCAGCTCCTGTACGAATTTACGACAAATAGCTCGTGACCTGAATAAATCTCTTCCTGAAT	543	
Db	480	CTCAGCTCCTGTACGAATTTACGACAAATAGCTCGTGACCTGAATAAATCTCTTCCTGAAT	539	
QY	544	ACTAATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	603	
Db	540	ACTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	599	
QY	604	AAAAAAA 610		
Db	600	AAAAAAA 606		

RESULT 3
B0619318

LOCUS DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANIS

BQ619318 RNOSEQ6D01_SK.ab1 salt stressed Zea mays cDNA clone RNOSEQ6D01_SK.ab1 similar to No homology, mRNA sequence. 648 bp mRNA linear EST 27-JUN-2002

BQ619318
BQ619318.1 GI:21621312

EST.
Zea mays

Zea mays

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 01:23:01 ; Search time 1439.61 Seconds
(without alignments)
7604.943 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 accaccgctcccccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_estl:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_mus:*
 - 24: em_gss_mam:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	508.6	75.2	648	14	BQ619167
2	508.6	75.2	648	14	BQ619315
3	508.6	75.2	648	14	BQ619318
4	508.6	75.2	648	14	BQ619337
5	508.6	75.2	648	14	BQ619383
6	508.6	75.2	648	14	BQ619390

7	498.2	73.7	523	10	AW331212
8	492.6	72.9	552	10	BE025302
9	481.2	71.2	524	10	AW288876
10	474.2	70.1	553	11	AY104409
11	470.6	69.6	500	10	AW787732
12	464.6	68.7	546	9	AA979839
13	462.8	68.5	550	12	BG840383
14	460	68.0	539	9	AI855425
15	456.4	67.5	496	10	BE19299
16	449	66.4	476	10	BE519259
17	449	66.4	481	9	AI964458
18	441.6	65.3	465	10	BE129897
19	429.6	63.6	436	10	AW787315
20	415.6	61.5	466	9	AI964534
21	409	60.5	443	10	BE129644
22	404.2	59.8	433	10	BE225008
23	361.2	53.4	401	10	AW289056
24	359	53.1	417	13	EM501439
25	323.6	47.9	541	10	AW745400
26	321.6	47.6	537	10	AW745400
27	319.6	47.3	326	10	BE025303
28	317	46.9	516	10	BE364814
29	316	46.7	519	10	AW680016
30	315.2	46.6	513	14	BQ280709
31	311.6	46.1	509	14	BQ280894
32	310.6	45.9	485	10	BE599123
33	310.6	45.9	522	10	AW679969
34	294.8	43.6	466	10	BE597738
35	275.4	40.7	362	12	BG840882
36	269	39.8	311	12	BF729420
37	247.2	36.6	399	13	BM318672
38	245.2	36.3	367	10	AW747603
39	226.8	33.6	348	10	BE593507
40	214.2	31.7	288	9	AI649804
41	212.6	31.4	274	10	AW288875
42	212	31.4	338	10	AW924579
43	185.4	27.4	512	14	BQ283209
44	185.4	27.4	583	14	BQ281558
45	183.6	27.2	277	10	BE593304

ALIGNMENTS

RESULT 1
BQ619167
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BQ619167
RNOSEQ4E05_SK.abi
mays cdna clone RNOSEQ4E05_SK.abi similar to No homology, mRNA
sequence.
BQ619167
EST.
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ4E05_SK.abi"

648 bp mRNA linear EST 27-JUN-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:52:33 ; Search time 29 Seconds
(without alignments)
331.081 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510
Sequence: 1 MAYTQEVYDCESEVRVAPAGFGRHGGVQVHVYKFEVDVTSRAGANHHHHHHGG 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	510	100.0	93	US-10-090-035-2	Sequence 2, Appli
2	510	100.0	93	US-10-090-035-4	Sequence 4, Appli
3	510	100.0	93	US-10-090-035-6	Sequence 6, Appli
4	507	99.4	93	US-10-090-035-10	Sequence 10, Appli
5	499.5	97.9	94	US-10-090-035-8	Sequence 8, Appli
6	499.5	97.9	94	US-10-090-035-16	Sequence 16, Appli
7	295	57.8	91	US-10-090-035-18	Sequence 18, Appli
8	277.5	54.4	91	US-10-090-035-22	Sequence 22, Appli
9	273	53.5	92	US-10-090-035-20	Sequence 20, Appli
10	258	50.6	92	US-10-090-035-24	Sequence 24, Appli
11	245	48.0	99	US-10-090-035-14	Sequence 14, Appli
12	80	15.7	359	US-10-232-563-7	Sequence 7, Appli
13	80	15.7	3352	US-09-888-615-64	Sequence 64, Appli
14	78	15.3	332	US-10-232-563-2	Sequence 2, Appli
15	77	15.1	359	US-10-232-563-6	Sequence 6, Appli
16	76	14.9	400	US-09-879-312-2	Sequence 2, Appli
17	73.5	14.4	285	US-10-278-173-16	Sequence 16, Appli
18	73.5	14.4	309	US-10-052-798-9	Sequence 9, Appli
19	73.5	14.4	312	US-10-052-798-10	Sequence 10, Appli

20	73	14.3	19	US-09-876-904A-627	Sequence 627, App
21	73	14.3	633	US-09-824-735-3	Sequence 3, Appli
22	73	14.3	633	US-09-801-368-338	Sequence 338, App
23	71.5	14.0	310	US-10-052-798-11	Sequence 11, Appli
24	71.5	14.0	354	US-09-825-302-501	Sequence 501, App
25	71.5	14.0	532	US-09-833-790-428	Sequence 428, App
26	69.5	13.6	315	US-10-232-563-11	Sequence 11, Appli
27	69.5	13.6	623	US-10-108-605-125	Sequence 125, App
28	69.5	13.6	623	US-10-108-605-129	Sequence 129, App
29	69.5	13.6	726	US-09-932-257A-19	Sequence 19, Appli
30	69	13.4	257	US-09-738-626-6135	Sequence 6135, Ap
31	68.5	13.4	120	US-09-051-013-3	Sequence 3, Appli
32	68	13.3	139	US-09-813-820-8	Sequence 8, Appli
33	67.5	13.2	118	US-09-205-658-120	Sequence 120, App
34	67.5	13.2	530	US-10-044-692-317	Sequence 317, App
35	67.5	13.2	530	US-10-044-539-317	Sequence 317, App
36	67.5	13.2	567	US-10-270-333-126	Sequence 126, App
37	67.5	13.2	605	US-09-741-233A-2	Sequence 2, Appli
38	67.5	13.2	1207	US-10-108-605-71	Sequence 71, Appli
39	67	13.1	446	US-09-853-386-69	Sequence 69, Appli
40	67	13.1	507	US-09-795-927-10	Sequence 10, Appli
41	67	13.1	507	US-09-738-897-2	Sequence 201, App
42	67	13.1	608	US-09-975-719-201	Sequence 200, App
43	67	13.1	639	US-09-975-719-200	Sequence 199, App
44	67	13.1	643	US-09-975-719-199	Sequence 70, Appli
45	67	13.1	643	US-09-853-386-70	

ALIGNMENTS

RESULT 1

US-10-090-035-2
; Sequence 2, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 357118/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

Query Match	100.0%;	Score 510;	DB 9;	Length 93;
Best Local Similarity	100.0%;	Pred. No. 5e+45;	Mismatches 0;	Indels 0; Gaps 0;
Matches 93;	Conservative			
QY	1	MAYTQEVYDCESEVRVAPAGFGRHGGVQVHVYKFEVDVTSRAGANHHHHHHHHGG	60	
DB	1	MAYTQEVYDCESEVRVAPAGFGRHGGVQVHVYKFEVDVTSRAGANHHHHHHHHHHGG	60	
QY	61	HGFVVRTRVEEDINTCTGCVHERRESFLARAN	93	
DB	61	HGFVVRTRVEEDINTCTGCVHERRESFLARAN	93	

RESULT 2

US-10-090-035-4
; Sequence 4, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 357118/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

us-10-090-035-2.rapb

Mon Jun 16 14:55:54 2003

```

; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272.227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-4

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHGGG 60
Db 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHGGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 3
US-10-090-035-6
; Sequence 6, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272.227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-6

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHGGG 60
Db 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHHHGGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 4
US-10-090-035-10
; Sequence 10, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272.227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-10

Query Match
Best Local Similarity 99.4%; Score 507; DB 9; Length 93;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHGGG 60
Db 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHHHGGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 5
US-10-090-035-8
; Sequence 8, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272.227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-8

Query Match
Best Local Similarity 97.9%; Score 499.5; DB 9; Length 94;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGAN-HHHHHHHGG 59
Db 1 MAYTQEVYDVCSEEVRSVAPAGFGRHGGVQGHVVKKEFEVDVTSRAGANHHHHHHHHHHGG 60
QY 60 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 6
US-10-090-035-18
; Sequence 18, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272.227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-18

```


; ORGANISM: Triticum aestivum
US-10-090-035-18

Query Match 97.9%; Score 499.5; DB 9; Length 94;
Best Local Similarity 98.9%; Pred. No. 5.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels

[illegible]

Qy	60	GHGFVVRETRVEEDINTCTGVHERRESFLARAN	93
Db	61	GHGFVVRETRVEEDINTCTGVHERRESFLARAN	94

RESULT 7

US-10-090-035-16
; Sequence 16, Application US/10090035
; Patent No. US20020170089A1

```

; GENERAL INFORMATION: Carl R.
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```

; LENGTH: 91

```

; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-090-035-16

```

Query Match 57.8%; Score 295; DB 9; Length 91;
Best Local Similarity 60.6%; Pred. No. 4e-23;
Matches 57; Conservative 14; Mismatches 19; Indels

QY	1 MAYYQEV	2 DYCSE	3 EYRSV	4 AP-	5 AGFGR	6 HGGV	7 QQHV	8 VKKE	9 FEE	10 VDT	11 VSR	12 ACAN	13 HHHH	14 HHG	15 HG	16 59
	:						:	:	:	:	:	:	:	:	:	:
Db	1 MAHYQEV	2 DYCSE <td>3 EYRSV</td> <td>4 TPTG</td> <td>5 GFLG</td> <td>6 RGVQ</td> <td>7 QQHV</td> <td>8 VKKE</td> <td>9 FEE</td> <td>10 VDT <td>11 VSR</td> <td>12 ACAN <td>13 HHHH <td>14 HHG <td>15 HG <td>16 57</td> </td></td></td></td></td>	3 EYRSV	4 TPTG	5 GFLG	6 RGVQ	7 QQHV	8 VKKE	9 FEE	10 VDT <td>11 VSR</td> <td>12 ACAN <td>13 HHHH <td>14 HHG <td>15 HG <td>16 57</td> </td></td></td></td>	11 VSR	12 ACAN <td>13 HHHH <td>14 HHG <td>15 HG <td>16 57</td> </td></td></td>	13 HHHH <td>14 HHG <td>15 HG <td>16 57</td> </td></td>	14 HHG <td>15 HG <td>16 57</td> </td>	15 HG <td>16 57</td>	16 57

```

QY      60  GHGFVRETRVEEDINTCTGEVHERRESFLARAN 93
          :|||:|||:|||:|||:|||:|||:|||:|||
Db      58  NDYLMVRETKEVEDNTCTGEPRKOSFLKSD 91

```

RESULT 8

US-10-090-035-22
; Sequence 22, Application US/10090035
; Patent No. US20020170089A1

```

; GENERAL INFORMATION: Carl R.
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22

```

: LENGTH: 91

US-10-090-035-22
; ORGANISM: Triticum aestivum
; TYPE: PRT

Query Match	54.48;	Score 277.5;	DB 9;	Length 91;
Best Local Similarity	61.28;	Pred. No. 2.4e-21;		

	Matches	60:	Conservative	12:	Mismatches	13:	Indels	13:	Gaps	6:
Qy	1	MAYIQEVDYCSSEVRSA-	PAGEGRGGGQQHVVKEF-	EVDVTYSRAGANHHHHGH	58					
.										
Dd	1	MAHFQDYDYSSEVRAVGNP-	-RRGGQGEHIVKETFVEFDTSGR----	RGHGHGH	53					
Qy	59	G---GHGFVYRETRVEDINTCTGEVHERRESFLARAN	93							
Dd	54	GRSGSH-FEYRESRLDEFTNRGEFVERKENFVRAD	90							

RESULT 9

US-10-090-035-20
; Sequence 20, Application US/10090035
; Patent No. US20020170089A1

```

: GENERAL INFORMATION: Carl R.
: APPLICANT: Simmons,
: TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
: TITLE OF INVENTION: Proteins and Uses Thereof
: FILE REFERENCE: 35718/242990
: CURRENT APPLICATION NUMBER: US/10/090,035
: CURRENT FILING DATE: 2002-02-28
: PRIOR APPLICATION NUMBER: 60/272,227
: PRIOR FILING DATE: 02/28/2001
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSEQ for Windows Version 4.0
: SEQ ID NO 20
: LENGTH: 92
: TYPE: PRT
: ORGANISM: Triticum aestivum
: US-10-090-035-20

```

Query Match

Query Match 53.5%; Score 273; DB 9; Length 92;
Best Local Similarity 61.6%; Pred. No. 7e-21;
Matches 61; Conservative 12; Mismatches 12; Indels

Qy 1 MAYQEVDYCSEEVRSVA - PAGEGRHG - GGVOQHVVKEF - EEVDTVSRAGANHHHHHGH 57
 ||::|||:|||||: || |||:||||: || | |||
 Db 1 MAHQEVDYCSEEVRAVGYP - --RGCGGQVEHIVETVQFEDTAGR - --RHGHGH 53

Qy	58	HG--GHGFVVRTRVEEDINTCGEVHERRESFLAR	93
	..	: : : : : : : :	
Db	54	HGRSGH-FEVRSRLEEDINTRTFGEFHEKKNFVRAD	91

RESULT 10

US-10-090-035-24
; Sequence 24, Application US/10090035
; Patent No. US20020170089A1

```

/ / 1. GENERAL INFORMATION:
/ / APPLICANT: Simmons, Carl R.
/ / TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
/ / TITLE OF INVENTION: Proteins and Uses Thereof
/ / FILE REFERENCE: 35718/242990
/ / CURRENT APPLICATION NUMBER: US/10/090,035
/ / CURRENT FILING DATE: 2002-02-28
/ / PRIOR APPLICATION NUMBER: 60/272,227
/ / PRIOR FILING DATE: 02/28/2001
/ / NUMBER OF SEQ ID NOS: 25
/ / SOFTWARE: FastSeq for Windows Version 4.0
/ / SEQ ID NO 24
/ / LENGTH: 92
/ / TYPE: PRT
/ / ORGANISM: Triticum aestivum
/ / US-10-090-035-24

```

Query Match

Query Match 50.6%; Score 258; DB 9; Length 92;
Best Local Similarity 58.6%; Pred. NO. 2.4e-19;
Matches 58; Conservative 14; Mismatches 13; Indels

QY 1 MAYQEVDCSEEVRSVA-PAGFGRHG-GGVQGHVYKEF-EVDTVSRAGANHHHHHGH 57
||:|||||: || | |||: ||| : || | |||
Db 1 MAHQEVDCSEEVRAVGYPA--FRGCGVQGHVYKETVQVEEDTAGR----RHGHGH 53

3

```

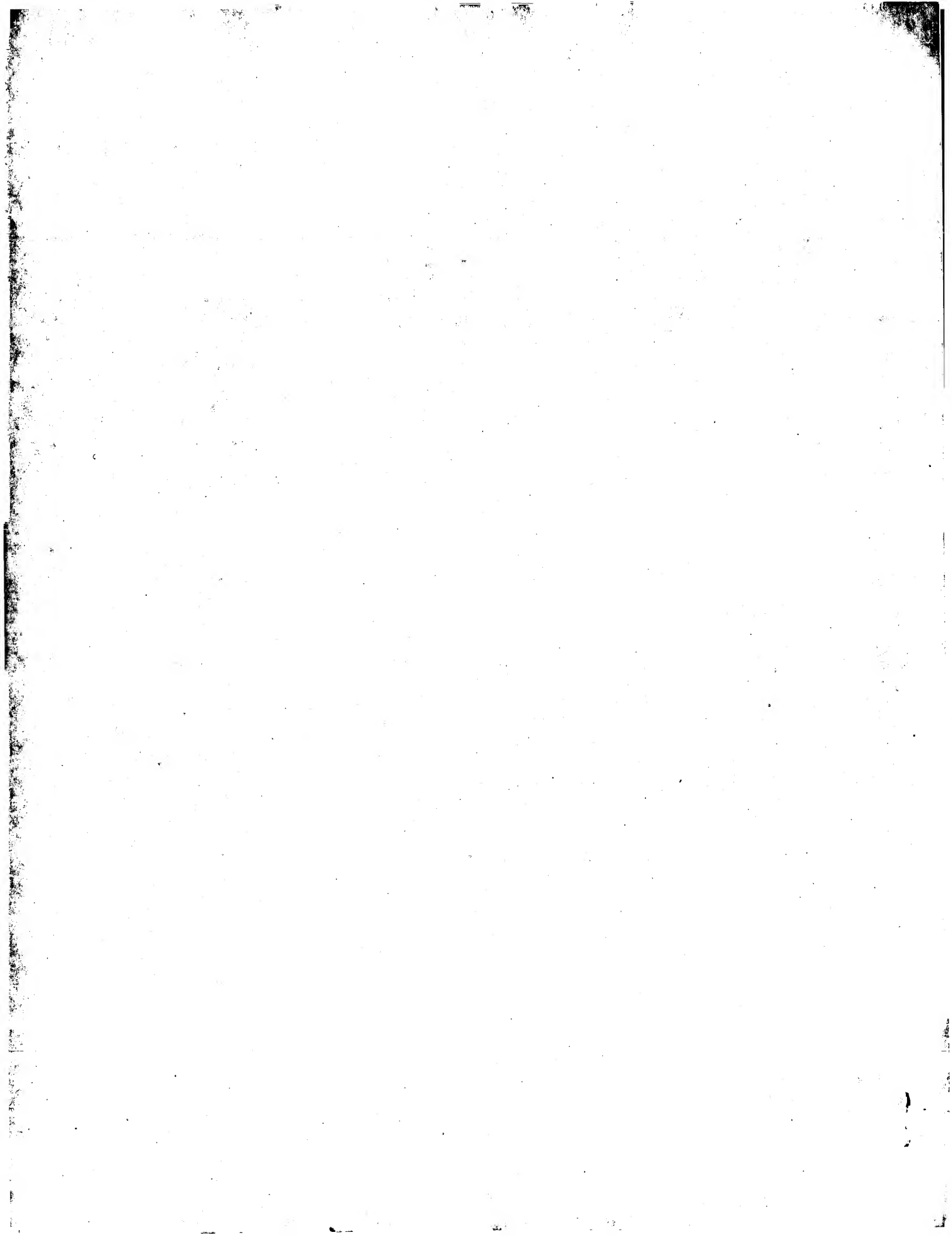
Db      230 QLVSMVRELNQLRGF 246
      :: |::: |
RESULT 15
US-10-232-563-6
; Sequence 6, Application US/10232563
; Publication No. US20030087394A1
; GENERAL INFORMATION:
; APPLICANT: Shatma, Arun
; TITLE OF INVENTION: INSULIN RELATED TRANSCRIPTION FACTOR AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 10276-072001
; CURRENT APPLICATION NUMBER: US/10/232,563
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/316,453
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-232-563-6

Query Match      15.1%; Score 77; DB 9; Length 359;
Best Local Similarity 25.0%; Pred. No. 2.8;
Matches 24; Conservative 7; Mismatches 29; Indels 36; Gaps 4;

QY      1 MAYQOEVDCSEEVRSVAP-----AGFGRHG--GVQQHVVKKEFEVDTVSRAGA-- 49
Db      117 MSGYQH--HLNPEALNLTPEDAVEALIGSGHHGHAHPAAAAAYEAFRGQSFAGGG 174
      :|::: |
QY      50 -----NHHHHHGHGGHG 62
      :|||| |
Db      175 ADDMGAGHHGCAHTAHHHSAHHHHHHHHHHGGSG 210
      :|||| |

```

Search completed: June 6, 2003, 13:02:10
 Job time : 30 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:20:30 ; Search time 39 Seconds
(without alignments)
731.306 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYQYDYDCSEVRVAPV.....INTCTGEVHERRESFLARAN 93

Scoring table:

	BLOSUM62	Xgapop 10, 0	Xgapext 0.5
Ygapop 10, 0			
Fgapop 6.0			
Delop 6.0			

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US10090035/runat_06062003_105504_10985/app_query.fasta_1.526
-DB=issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10090035 -ECGN=1.1.55 -runat_06062003_105504_10985 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	83	16.3	4403765	4	US-09-103-840A-2
C 2	83	16.3	4411529	4	US-09-103-840A-1
C 3	80.5	15.8	2353	5	PCT-US92-06840-1
C 4	80	15.7	2310	3	US-09-036-987A-25
C 5	80	15.7	2310	4	US-09-370-700-25
C 6	79.5	15.6	195	2	US-08-622-740-9
C 7	79.5	15.6	195	3	US-08-440-689-9
C 8	79.5	15.6	195	4	US-09-122-399-9
9	77	15.1	3218	2	US-08-789-275-1
10	77	15.1	3218	2	US-08-677-862-1
11	77	15.1	3218	2	US-09-252-571-1
12	77	15.1	3218	3	US-09-434-065-1

13	76	14.9	1203	4	US-09-086-010-1	Sequence 1, Appl
14	74	14.5	1305	4	US-09-329-234A-6	Sequence 6, Appl
C 15	74	14.5	2121	1	US-08-240-783B-1	Sequence 1, Appl
C 16	74	14.5	2121	3	US-09-084-813-1	Sequence 1, Appl
C 17	74	14.5	2121	5	PCT-US92-09662-1	Sequence 38, Appl
C 18	73.5	14.4	123	2	US-08-273-146-38	Sequence 14, Appl
C 19	73.5	14.4	185	4	US-08-447-985-14	Sequence 6, Appl
C 20	73.5	14.4	930	4	US-09-079-029-6	Sequence 7, Appl
C 21	73.5	14.4	939	4	US-09-079-029-7	Sequence 11, Appl
C 22	73.5	14.4	3120	1	US-08-491-146-11	Sequence 1, Appl
C 23	73.5	14.4	3120	1	US-08-234-011-1	Sequence 9, Appl
C 24	73.5	14.4	3120	1	US-08-701-062A-1	Sequence 3, Appl
C 25	73.5	14.4	3120	2	US-08-072-064-7	Sequence 2, Appl
C 26	73	14.3	1970	3	US-08-072-064-7	Sequence 5, Appl
C 27	73	14.3	1970	5	PCT-US92-08558-3	Sequence 1, Appl
C 28	73	14.3	2059	5	PCT-US92-08558-2	Sequence 2, Appl
C 29	73	14.3	2066	3	US-08-072-064-3	Sequence 3, Appl
C 30	73	14.3	2066	3	US-08-072-064-3	Sequence 5, Appl
C 31	73	14.3	2066	3	US-08-072-064-3	Sequence 7, Appl
C 32	73	14.3	2066	3	US-08-072-064-3	Sequence 1, Appl
C 33	72.5	14.2	377	2	US-08-332-766A-1	Sequence 1, Appl
C 34	72.5	14.2	1723	1	US-08-241-766-1	Sequence 2, Appl
C 35	72.5	14.2	1723	1	US-08-241-766-2	Sequence 8, Appl
C 36	72.5	14.2	2232	1	US-08-241-766-12	Sequence 3, Appl
C 37	71.5	14.0	933	4	US-09-079-029-8	Sequence 5, Appl
C 38	71.5	14.0	4131	1	US-08-309-512-3	Sequence 96, Appl
C 39	71.5	14.0	4131	5	PCT-US92-08756A-3	Sequence 2, Appl
C 40	71.5	14.0	4392	1	US-08-026-138E-5	Sequence 1, Appl
C 41	71	13.9	274	2	US-08-448-418-96	Sequence 1, Appl
C 42	71	13.9	71989	4	US-09-443-501A-2	Sequence 7, Appl
C 43	70.5	13.8	883	3	US-09-184-658-7	Sequence 1, Appl
C 44	70.5	13.8	43280	2	US-08-804-227C-1	Sequence 2, Appl
C 45	70.5	13.8	4403765	4	US-09-103-840A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.22e+04
Score: 83.00
Percent Similarity: 46.55%
Best Local Similarity: 36.21%
Query Match: 16.27%
DB: 4
Length: 4403765
Matches: 21
Conservative: 6
Mismatches: 29
Indels: 2
Gaps: 1

US-10-090-035-2 (1-93) x US-09-103-840A-2 (1-4403765)

[illegible]

Db 483 GCCCATGATCGCTCTGCAGCGCTGTCACCGAGGAC-----CCGACCGAGGT 530
QY 81 LHisGluArgArgGluSerPheLeuAlaArg 91
Db 531 GCACCAACCAGGAGGTGATCTGCTGTCGAG 561

RESULT 4

US-09-036-987A-25/c
; Sequence 25, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: .Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036.987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

MOLECULE TYPE: DNA (genomic)

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..1077
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1165..1992

US-09-036-987A-25

Alignment Scores:

Pred. No.: 1.24 Length: 2310
Score: 80.00 Matches: 27
Percent Similarity: 41.76% Conservative: 11
Best Local Similarity: 29.67% Mismatches: 35
Query Match: 15.69% Indels: 18
DB: 3 Gaps: 4

US-10-090-035-2 (1-93) x US-09-036-987A-25 (1-2310)

QY 7 ValAspTyrCysSerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGly 26
Db 480 GTAGACCTCGTGGTGGAAACGTGCACGAACTTGCAGTCTGCCGTCGAGCGC----- 427

QY 27 GlyGlyValGlnGlnHisValValLysGluLysPheGluValAlaAspThrValSerArg 46
Db 426 ---GGCCTGTCAGCAGCAGCTTGGTGGCGGACACCGTGGTGTATCACAAGCGCTC---CGA 373
QY 47 AlaGly-----AlaAsnHisHisHisHisHisHis 55
Db 372 GCCGGTGTATCAGCGGTGCGAGCGGTGGTGGCGGCGCAAGTGCACCAACCGTCCACGCC 313
QY 56 GlyHisHisGlyGlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsn 75
Db 312 GGACATCAGCGCGCAACACAGTTCGCGGTGCGAGATGTCGCCCGCAGACGAA----- 262
QY 76 ThrCysThrGlyGluValHisGluArgGlu 86
Db 261 CTTACGCGCGGGGTGTTCGCGGACCGCGCCAG 229

RESULT 5

US-09-370-700-25/c
; Sequence 25, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:

; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide

; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 25

LENGTH: 2310

TYPE: DNA

ORGANISM: Saccharopolyspora spinosa

FEATURE:

NAME/KEY: CDS

LOCATION: (88)..(1077)

FEATURE:

NAME/KEY: CDS

LOCATION: (1165)..(1992)

US-09-370-700-25

Alignment Scores:

Pred. No.: 1.24 Length: 2310
Score: 80.00 Matches: 27
Percent Similarity: 41.76% Conservative: 11
Best Local Similarity: 29.67% Mismatches: 35
Query Match: 15.69% Indels: 18
DB: 4 Gaps: 4

US-10-090-035-2 (1-93) x US-09-370-700-25 (1-2310)

QY 7 ValAspTyrCysSerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGly 26
Db 480 GTAGACCTCGTGGTGGAAACGTGCACGAACTTGCAGTCTGCCGTCGAGCGC----- 427
QY 27 GlyGlyValGlnGlnHisValValLysGluLysPheGluValAlaAspThrValSerArg 46
Db 426 ---GGCCTGTCAGCAGCAGCTTGGTGGCGGACACCGTGGTGTATCACAAGCGCTC---CGA 373
QY 47 AlaGly-----AlaAsnHisHisHisHisHisHis 55
Db 372 GCCGGTGTATCAGCGGTGCGAGCGGTGGTGGCGGCGCAAGTGCACCAACCGTCCACGCC 313
QY 56 GlyHisHisGlyGlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsn 75

Sequence 9, Application US/08440689
Patent No. 6025545
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
TITLE OF INVENTION: and Cells Thereof
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/440,689
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.0130S2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-440-689-9

Alignment Scores:
Pred. No.: 0.0531 Length: 195
Score: 79.50 Matches: 23
Percent Similarity: 48.98% Conservative: 1
Best Local Similarity: 46.94% Mismatches: 13
Query Match: 15.59% Indels: 12
DB: 3 Gaps: 3

US-10-090-035-2 (1-93) x US-08-440-689-9 (1-195)

QY 16 SerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLys 35
Db 153 TCTGGAGGACCGCGGGCGGCGGAGGCTGGCGTGGACTT-----112
QY 36 GluLysPheGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHis 55
Db 111 GAGCCCTCGAAGCGGAGCGGCGGCTGGCGGAGGC-----CATCATCAC 64

QY 56 Gly-HisHisGlyGly---HisGly 62
Db 63 GGTTCATCATCGGTGGCGCCATGGC 39

RESULT 8
US-09-122-399-9/c
Sequence 9, Application US/09122399
Patent No. 6329574
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants

312 GGACATCAGCGCGCAACACAGTTCGCGTTCGAGATGTCGCGCGACAGCAA-----262
QY 76 ThrCysThrGlyGluValHisGluArgArgGlu 86
Db 261 CTTACGCGGGGTGTCTCCGCGACCGCGCCAG 229

RESULT 6
US-08-622-740-9/c
Sequence 9, Application US/08622740
Patent No. 590390
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
TITLE OF INVENTION: and Cells Thereof
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/622,740
FILING DATE: 27-MARCH-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.0130S4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-622-740-9

Alignment Scores:
Pred. No.: 0.0531 Length: 195
Score: 79.50 Matches: 23
Percent Similarity: 48.98% Conservative: 1
Best Local Similarity: 46.94% Mismatches: 13
Query Match: 15.59% Indels: 12
DB: 3 Gaps: 3

US-10-090-035-2 (1-93) x US-08-622-740-9 (1-195)

QY 16 SerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLys 35
Db 153 TCTGGAGGACCGCGGGCGGCGGAGGCTGGCGTGGACTT-----112
QY 36 GluLysPheGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHis 55
Db 111 GAGCCCTCGAAGCGGAGCGGCGGCTGGCGGAGGC-----CATCATCAC 64

QY 56 Gly-HisHisGlyGly---HisGly 62
Db 63 GGTTCATCATCGGTGGCGCCATGGC 39

RESULT 7
US-08-440-689-9/c

TITLE OF INVENTION: and Cells Thereof
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schweigman, Lundberg & Woessner, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/122,399
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/112,245
FILING DATE: 25-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.13US01
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-122-399-9

Alignment Scores:
Pred. No.: 0.0531 Length: 195
Score: 79.50 Matches: 23
Percent Similarity: 48.98% Conservative: 1
Best Local Similarity: 46.94% Mismatches: 13
Query Match: 15.59% Indels: 12
DB: 4 Gaps: 3

US-10-090-035-2 (1-93) x US-09-122-399-9 (1-195)

QY 16 SerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLys 35
Db 153 TCTGGAGGACGGCGGGGAGCGGTGGCTGGCTT-----112
QY 36 GluLysPheGluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHis 55
Db 111 GAGCCCTGGACGCGGCGGCGGTGGCGGAGGC-----CATCATCAC 64
QY 56 Gly-HisHisGlyGly---HisGly 62
Db 63 GGTTCATCACGGTGGCGGCATGGC 39

RESULT 9
US-08-789-275-1
Sequence 1, Application US/08789275A
Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Palleja, Estivill Xavier
APPLICANT: Pritchard, Melanie.
APPLICANT: Vilario, Jordi Guilmera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the
TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
FILE REFERENCE: U 01114-4
CURRENT APPLICATION NUMBER: US/08/789,275A

CURRENT FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 1
LENGTH: 2373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-08-789-275-1

Alignment Scores:
Pred. No.: 3.05 Length: 2373
Score: 77.00 Matches: 22
Percent Similarity: 39.39% Conservative: 4
Best Local Similarity: 33.33% Mismatches: 26
Query Match: 15.10% Indels: 14
DB: 4 Gaps: 2

US-10-090-035-2 (1-93) x US-08-789-275-1 (1-2373)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly-----27
Db 1738 AGTCCCGAGGTGGCTCAGCAATTTCTGCTCTCTTGGTTGGTCAGGCACCTGAAGCTCCT 1797
QY 28 -----GlyValGlnGlnHisValLysGluLysPheGluValAspThrVal 44
Db 1798 ACACAGGTCACTGTTGAAACTCATCTGTTCAAGAAACAACTTTTCATGTAGCCCTCAA 1857
QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisHisHisHisHisHis 56
Db 1858 CAGATGCAATGTCATCATCACCATGTAACAGTTCCCATCATCACCCACCACCACCAC 1917
QY 57 HisHisGlyGlyHisGly 62
Db 1918 CATCACCCACCACCATGGA 1935

RESULT 10

US-08-677-862-1
Sequence 1, Application US/08677862
Patent No. 5874230
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTHE, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,862
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: 796-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3218 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

US-10-090-035-2 (1-93) x US-09-252-571-1 (1-3218)

QY 11 SerGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
DB 1783 AGTCCCAGGTGCTCAGCAATTCCTGCTCCTCTGTTGGTTCAGGCACTGAAGCTCCT 1842
QY 28 -----GlyValGlnHisValValysGluLysPheGluValAlaAspThrVal 44
DB 1843 ACACAGGTCACTGTTGAACATCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisGly 56
DB 1903 CAGAAATGATGTCATCATCACCATTGTAACAGTTCCCATCACCACCAACCAC 1962

QY 57 HisHisGlyGlyHisGly 62
DB 1963 CATCACCACCACCATGGA 1980

RESULT 12
US-09-434-065-1
; Sequence 1, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/434,065
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-434-065-1

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

US-10-090-035-2 (1-93) x US-08-677-862-1 (1-3218)

QY 11 SerGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
DB 1783 AGTCCCAGGTGCTCAGCAATTCCTGCTCCTCTGTTGGTTCAGGCACTGAAGCTCCT 1842
QY 28 -----GlyValGlnHisValValysGluLysPheGluValAlaAspThrVal 44
DB 1843 ACACAGGTCACTGTTGAACATCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisGly 56
DB 1903 CAGAAATGATGTCATCATCACCATTGTAACAGTTCCCATCACCACCAACCAC 1962

QY 57 HisHisGlyGlyHisGly 62
DB 1963 CATCACCACCACCATGGA 1980

RESULT 11
US-09-252-571-1
; Sequence 1, Application US/09252571
; Patent No. 5981250
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/252,571
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-252-571-1

Percent Similarity: 39.39% Conservative: 4
 Best Local Similarity: 33.33% Mismatches: 26
 Query Match: 15.10% Indels: 14
 DB: 3 Gaps: 2

US-10-090-035-2 (1-93) x US-09-434-065-1 (1-3218)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
 DB 1783 ACTCCAGGTGTCGACCAATTCCTCTCTTGGTGGTCAGGCACTGAAGCTCCT 1842
 QY 28 -----GlyValGlnGlnHisValValLysGluLysPheGluGluValAspThrVal 44
 DB 1843 ACACAGGTCACTGTTGAACATCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
 QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisHisHisGly 56
 DB 1903 CAGATGTCATGTCATCATCCAGTGTAAAGTTCCTCATCACCACCACTCACCACCA 1962
 QY 57 HisHisGlyGlyHisGly 62
 DB 1963 CATCACCACCACTGGA 1980

RESULT 13

US-09-086-010-1
 ; Sequence 1, Application US/09086010
 ; Patent No. 6274338
 ; GENERAL INFORMATION:
 ; APPLICANT: Glmcher, Laurie H. et al.
 ; TITLE OF INVENTION: Human c-Maf Compositions and
 ; TITLE OF INVENTION: Methods of Use Thereof
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA: HUI-027CP
 ; APPLICATION NUMBER: US/09/086,010
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/030,579
 ; FILING DATE: 2-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kara, Catherine J.
 ; REGISTRATION NUMBER: 41,106
 ; REFERENCE/DOCKET NUMBER: HUI-027CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)42-4214
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1203 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1203
 ; US-09-086-010-1

Alignment Scores:
 Pred. No.: 1.65 Length: 1203
 Score: 76.00 Matches: 18

Percent Similarity: 45.61% Conservative: 8
 Best Local Similarity: 31.58% Mismatches: 25
 Query Match: 14.90% Indels: 6
 DB: 4 Gaps: 1

US-10-090-035-2 (1-93) x US-09-086-010-1 (1-1203)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGlyValGln 30
 DB 403 GCCACGACGTAGCCGCCGCCGCCGCCGCCGCCCTCTTGGCGCGCACGCCGC 462
 QY 31 GlnHis-----ValValLysGluLysPheGluGluValAspThrVal 44
 DB 463 GAGGAGATGGCGCCGCCGCCGCCGCCGCCGCCGTGTCGCCGTGATCGCGCGCGCGCGCG 522
 QY 45 SerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyHisGlyHis 61
 DB 523 AGCGGCG 573

RESULT 14

US-09-329-234A-6
 ; Sequence 6, Application US/09329234A
 ; Patent No. 6331416
 ; GENERAL INFORMATION:
 ; APPLICANT: ShanL, ziv
 ; APPLICANT: Shoseyov, Oded
 ; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND R
 ; TITLE OF INVENTION: PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTUR
 ; FILE REFERENCE: 00/20274
 ; CURRENT APPLICATION NUMBER: US/09/329,234A
 ; CURRENT FILING DATE: 2001-09-04
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 1305
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Construct containing Protein L, fused to CBDcex sequence
 ; US-09-329-234A-6

Alignment Scores:

Pred. No.: 3.27 Length: 1305
 Score: 74.00 Matches: 16
 Percent Similarity: 40.91% Conservative: 2
 Best Local Similarity: 36.36% Mismatches: 22
 Query Match: 14.51% Indels: 4
 DB: 4 Gaps: 1

US-10-090-035-2 (1-93) x US-09-329-234A-6 (1-1305)

QY 19 Proalagly-----PheGlyArgHisGlyGlyValGlnHisValVal 34
 DB 1173 CCG 1232
 QY 35 LysGluLysPheGluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHis 54
 DB 1233 CCGACGCGGTTCGCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1292
 QY 55 HisGlyHisHis 58
 DB 1293 CACCAACCAAC 1304

RESULT 15

US-08-240-783B-1/c
 ; Sequence 1, Application US/08240783B
 ; Patent No. 5756348
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Kelli
 ; APPLICANT: Borden, Laurence A.
 ; APPLICANT: Branchek, Theresa
 ; APPLICANT: Hartig, Paul R.
 ; APPLICANT: Weinshank, Richard L.

;; TITLE OF INVENTION: DNA ENCODING A GLYCINE TRANSPORTER AND USES
;; TITLE OF INVENTION: THEREOF
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham LLP
;; STREET: 1185 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.24
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/240,783B
;; FILING DATE:
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: White, John P.
;; REGISTRATION NUMBER: 28,678
;; REFERENCE/DOCKET NUMBER: 1795/39875-A-PCT-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 278-0400
;; TELEFAX: (212) 391-0525
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2121 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: unknown
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: N
;; ANTI-SENSE: N
;; ORIGINAL SOURCE:
;; ORGANISM: RAT GLYCINE TRANSPORTER
;; CELL TYPE: MAMMALIAN
;; CELL LINE: COS7
;; IMMEDIATE SOURCE:
;; CLONE: IB20a
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 62..1975
;; OTHER INFORMATION:
;;
US-08-240-783B-1

Alignment Scores:
Pred. No.: 6.25 Length: 2121
Score: 74.00 Matches: 24
Percent Similarity: 38.10% Conservatives: 8
Best Local Similarity: 28.57% Mismatches: 38
Query Match: 14.51% Indels: 14
DB: 3
Gaps: 3

US-10-090-035-2 (1-93) x US-08-240-783B-1 (1-2121)

QY	10	CysSerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGly-----	26
Db	535	TGTGTTCCAGGATTATTGCAGTAAGCCAGGCGACATGCGTCATGGACGAGAAGAA	476
QY	27	-----GlyGlyValGlnHisValValLysGluLysPheGluGluValAspThr	43
Db	475	GTAGTAGAAGCGATGCAGATGACACCGTTGTA-----GTAGATACC	434
QY	44	ValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisHisGlyHisGlyPhe	63
Db	433	GATGTACGT-----GGACACCAACCATCATACCATGCCCGCTTTGAACATGGGCT	380
QY	64	ValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGluValHisGlu	83
Db	379	GATCCTCCAGACCCCGAGCGCCCTTGCTTGCACAACTGGCCGAGGAGAGCTCCATGAA	320

QY 84 ArgArgGluSer 87
Db 319 GAAGAGAGGAAT 308
Search completed: June 16, 2003, 12:40:53
Job time : 1414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 11:41:43 ; Search time 125 Seconds
(without alignments)
1077.357 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYQYDYCSEVRVAPV.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xmlh
-MODEL-frame-p2n.model -USPTO.spool/US10090035/runat_06062003_105507_11116/app-query.fasta_1.526
-DB-publishedApplications NA -OFMT-fastap -SUFFIX-rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -SPARM-1 -END-1 -MATRIX-blosum62
-TRANS-human40.cdi -LIST-45 -DOALIGN-200 -THR_SCORE-pct -THR_MAX=100
-THR_MIN=0 -ALIGN-15 -MODS-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER-US10090035@cgn_1_1176 @runat_06062003_105507_11116
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	574	9	US-10-090-035-3
2	510	100.0	577	9	US-10-090-035-5
3	510	100.0	676	9	US-10-090-035-1
4	507	99.4	529	9	US-10-090-035-9
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 9, Appli

5 499.5 97.9 524 9 US-10-090-035-17
6 499.5 97.9 580 9 US-10-090-035-7
7 295 54.4 436 9 US-10-090-035-15
8 277.5 54.4 436 9 US-10-090-035-21
9 258 50.6 584 9 US-10-090-035-19
10 258 50.6 584 9 US-10-090-035-23
11 251 49.2 348 9 US-10-090-035-13
12 213 41.8 237 10 US-09-923-876-2788
13 92.5 18.1 328 10 US-09-924-035A-17
14 92.5 18.1 591 9 US-09-938-842A-1858
15 88.5 17.4 546 10 US-09-728-445-817
16 83 16.3 3870 9 US-09-712-363-113
17 82.5 16.2 3966 9 US-10-198-846-13254
18 80 15.7 10062 10 US-09-888-615-5
19 78.5 15.4 1602 10 US-09-833-790-417
20 78 15.3 1059 9 US-10-232-563-4
21 78 15.3 1189 10 US-09-900-237-5
22 78 15.3 2681 9 US-10-232-563-1
23 77 15.1 1080 9 US-10-232-563-5
24 77 15.1 6381 10 US-09-969-347-216
25 76 14.9 1203 9 US-09-879-312-1
26 76 14.9 2627 9 US-10-037-270-434
27 75.5 14.8 491 9 US-09-854-133-343
28 75.5 14.8 491 10 US-09-738-973-343
29 75.5 14.8 25309 9 US-10-291-737-3
30 75 14.7 8979 9 US-09-738-626-2739
31 75 14.7 3309400 9 US-09-738-626-1
32 74.5 14.6 1458 9 US-10-103-313-83
33 74.5 14.6 1519 9 US-10-153-668-95
34 74 14.5 484 10 US-09-833-790-30
35 74 14.5 558 9 US-09-998-660-1
36 74 14.5 593 10 US-09-864-761-6743
37 74 14.5 910 12 US-10-062-254-325
38 74 14.5 32082 9 US-09-764-891-9679
39 73.5 14.4 279 9 US-09-984-130-143
40 73.5 14.4 858 9 US-09-938-842A-1494
41 73.5 14.4 930 12 US-10-052-798-6
42 73.5 14.4 939 12 US-10-052-798-7
43 73.5 14.4 1084 9 US-10-278-173-15
44 73.5 14.4 2331 9 US-09-938-842A-457
45 73 14.3 472 9 US-09-918-995-6480

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: SIMMONS, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Alignment Scores: 1.3e-58 Length: 574
Pred. No.: 510.00 Matches: 93
Score: 510.00
Percent Similarity: 100.00% Conservative: 0

Mon Jun 16 14:55:54 2003

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-3 (1-574)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProLa 20
DB 96 ATGGCTTACTACAGGAGGTGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 155
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 156 GGCTTGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCAAGGAGAAGTTCGAGGAG 215
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 216 GTGACACGGTCTCACGCGCGCGCGCCACACACCATGGTCCACACGCGCGC 275
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 276 CACGGCTTCGTGGCGCGAGACGAGGTGCGAAGAGGACATCAACACCTGCACCGCGGAG 335
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 336 GTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 374

RESULT 2

US-10-090-035-5
; Sequence 5, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(377)
US-10-090-035-5

Alignment Scores:
Pred. No.: 1,31e-58 Length: 577
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-5 (1-577)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProLa 20
DB 99 ATGGCTTACTACAGGAGGTGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 158
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 159 GGCTTGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCAAGGAGAAGTTCGAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 219 GTGACACGGTCTCACGCGCGCGCGCCACACCATGGTCCACACGCGCGC 278
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80

DB 279 CACGGCTTCGTGGTGGCGGAGACGAGGTGCGAAGAGGACATCAACACCTGCACCGCGGAG 338
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 339 GTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 377

RESULT 3

US-10-090-035-1
; Sequence 1, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(367)
US-10-090-035-1

Alignment Scores:
Pred. No.: 1,59e-58 Length: 676
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-1 (1-676)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProLa 20
DB 89 ATGGCTTACTACAGGAGGTGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 148
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 149 GGCTTGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCAAGGAGAAGTTCGAGGAG 208
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 209 GTCCACACGGTATCACGCGCGCGCGCCACACCATGGTCCACCGCGCGC 268
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 269 CACGGCTTCGTGGTGGCGGAGACGAGGTGCGGAGGAGACATCAACACCTGCACCGCGGAG 328
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 329 GTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 367

RESULT 4

US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227

; PRIOR FILING DATE: 02/28/2001
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 529
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (53)...(331)
 US-10-090-035-9

Alignment Scores:
 Pred. No.: 2,93e-58 Length: 529
 Score: 507.00 Matches: 92
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 99.41% Indels: 0
 DB: Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-9 (1-529)

QY 1 MetAlaTyrrGlnGluValAlaAspTyrCysSerGluGluValAlaArgSerValAlaProAla 20
 Db 53 ATGGCTTACTACAGAGAGTGGACTACTGCTCGAGGAGGTGAGGTGCTGGCCCGGCC 112
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluPheGluGlu 40
 Db 113 GGCTTCGGCCGCCAGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 172
 QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisHisHis 60
 Db 173 GTCGACACGGTCTCCAGCGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 172
 QY 61 HisGlyPheValValAlaArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
 Db 233 CACGCTCTGCTGGTGGCGGAGCAGGAGGTGAGGAGGTGAGGAGGTGAGGAGGTGAGGAGGTGAGGAG 292
 QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 Db 293 GTCCACGAGCGCAGGAGAGGCTTCTCGCCAGGCGCTAAC 331

RESULT 5
 US-10-090-035-17
 ; Sequence 17, Application US/10090035
 ; Patent No. US20020170089A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Simmons, Carl R.
 ; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
 ; FILE REFERENCE: 35718/242990
 ; CURRENT APPLICATION NUMBER: US/10/090, 035
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: 60/272,227
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 524
 ; TYPE: DNA
 ; ORGANISM: Triticum aestivum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (57)...(338)
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(524)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-090-035-17

Alignment Scores:
 Pred. No.: 2,85e-57 Length: 524
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0

Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: Gaps: 1

US-10-090-035-2 (1-93) x US-10-090-035-17 (1-524)

QY 1 MetAlaTyrrGlnGluValAlaAspTyrCysSerGluGluValAlaArgSerValAlaProAla 20
 Db 57 ATGGCTTACTACAGAGAGTGGACTACTGCTCGAGGAGGTGAGGTGCTGGCCCGGCC 116
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluPheGluGlu 40
 Db 117 GGCTTCGGCCGCCAGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 176
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn--HisHisHisHisHisHisHisHisHisHis 59
 Db 177 GTCGACACGGTCTCCAGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 236
 QY 60 GlyHisGlyPheValValAlaArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
 Db 237 GGCCACGGCTTCTGGTGGCGGAGCAGGAGGTGAGGAGGTGAGGAGGTGAGGAGGTGAGGAG 296
 QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 Db 297 GAGGTCCAGCGCAGGAGAGGCTTCTCGCCAGGCGCTAAC 338

RESULT 6

US-10-090-035-7
 ; Sequence 7, Application US/10090035
 ; Patent No. US20020170089A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Simmons, Carl R.
 ; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
 ; FILE REFERENCE: 35718/242990
 ; CURRENT APPLICATION NUMBER: US/10/090, 035
 ; CURRENT FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: 60/272,227
 ; PRIOR FILING DATE: 02/28/2001
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 580
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (99)...(380)
 US-10-090-035-7

Alignment Scores:
 Pred. No.: 3,24e-57 Length: 580
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: Gaps: 1

US-10-090-035-2 (1-93) x US-10-090-035-7 (1-580)

QY 1 MetAlaTyrrGlnGluValAlaAspTyrCysSerGluGluValAlaArgSerValAlaProAla 20
 Db 99 ATGGCTTACTACAGAGAGTGGACTACTGCTCGAGGAGGTGAGGTGCTGGCCCGGCC 158
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluPheGluGlu 40
 Db 159 GGCTTCGGCCGCCAGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 218
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn--HisHisHisHisHisHisHisHisHisHis 59
 Db 219 GTCGACACGGTCTCCAGCGCGCGGCTCCAGGAGGTGAGGTGCTCAAGGAGAGTTTCGAGGAG 278
 QY 60 GlyHisGlyPheValValAlaArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79

```
Db      279  GGCCACGGCTTCGTGTCGGCAGACACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 338
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      339  GAGGTCACGAGCGGAGGAGAGCTTCCTCGCCAGGGGCTAAC 380

RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.: 4e-30      Length: 591
Score: 295.00      Matches: 57
Percent Similarity: 75.53%      Conservative: 14
Best Local Similarity: 60.64%      Mismatches: 19
Query Match: 57.84%      Indels: 4
DB: 9      Gaps: 2
US-10-090-035-2 (1-93) x US-10-090-035-15 (1-591)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db      61  ATGGCTACTACACAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGCACCCACC 120
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db      121  GGCGGCTTCCTCGCGCGCGCGGCGGTGCAGCAGCAGCAGCTCGTCAAGGAGAGCGTTCCAG 180
QY      40  GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db      181  GAGATCGAC-----AGTCCGGCTCCGGCGCCGCCACACACACACACACCGCCG 231
QY      60  GlyHisGlyPheValArgGluThrArgValGluGluAspLeuAlaAsnThrCysThrGly 79
Db      232  AACGACTACCTGATGGTGGCGGAGACCAAGGTGGAGGAGGACTTCAACACCTGCACCGGC 291
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      292  GAGTTCGCGGAGCGGAGCAGAGCTTCCTGCTCAAGTCCGAC 333

RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:
Pred. No.: 5.68e-28      Length: 436
Score: 277.50      Matches: 60
Percent Similarity: 73.47%      Conservative: 12
Best Local Similarity: 61.22%      Mismatches: 13
Query Match: 54.41%      Indels: 13
DB: 9      Gaps: 6
US-10-090-035-2 (1-93) x US-10-090-035-21 (1-436)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db      54  ATGGCGCACTTCAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGCGGTGGCCACCG 113
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPhe--- 38
Db      114  GCC-----CGCCGCGCGCGCGGTGCAGGAGACATCGTCAAGGAGCGTTCGTG 164
QY      39  GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHis 58
Db      165  CAGGAGTTCGACACCTCCGCGCGC-----CGCCACGGTCACCGGTCCACCAC 212
QY      59  Gly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspLeu 75
Db      213  GGCGCGCGCTCGGTGCAC---TTCGAGGTGCGGAGAGCAGGCTCGAGGAGACTTCAAC 269
QY      76  ThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      270  ACCCGCAGCGGGGAGTTCACGAGCGCAAGGAGAACTTCGTGCTCAGGCGCGAT 323

RESULT 9
US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:
```

```
Db      279  GGCCACGGCTTCGTGTCGGCAGACACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 338
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      339  GAGGTCACGAGCGGAGGAGAGCTTCCTCGCCAGGGGCTAAC 380

RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.: 4e-30      Length: 591
Score: 295.00      Matches: 57
Percent Similarity: 75.53%      Conservative: 14
Best Local Similarity: 60.64%      Mismatches: 19
Query Match: 57.84%      Indels: 4
DB: 9      Gaps: 2
US-10-090-035-2 (1-93) x US-10-090-035-15 (1-591)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db      61  ATGGCTACTACACAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGCACCCACC 120
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db      121  GGCGGCTTCCTCGCGCGCGCGGCGGTGCAGCAGCAGCAGCTCGTCAAGGAGAGCGTTCCAG 180
QY      40  GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db      181  GAGATCGAC-----AGTCCGGCTCCGGCGCCGCCACACACACACACCGCCG 231
QY      60  GlyHisGlyPheValArgGluThrArgValGluGluAspLeuAlaAsnThrCysThrGly 79
Db      232  AACGACTACCTGATGGTGGCGGAGACCAAGGTGGAGGAGGACTTCAACACCTGCACCGGC 291
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      292  GAGTTCGCGGAGCGGAGCAGAGCTTCCTGCTCAAGTCCGAC 333

RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
```


Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-2 (1-93) x US-10-090-035-19 (1-584)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 46 ATGGCGCAGTTCAGGAGGTGGACTACTCTCGAGGAGTTCAGGCGGTGGCTACCG 105
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
Db 106 GCC-----CGCGCGCGTTCGCGCGGTTCGAGGAGACATCTCAAGGAGAGCGTTC 156
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHis 57
Db 157 GTGCAGGAGTTCGACACCGCGCGCGC-----CGCCAYGGTCCACCGGTTCAC 204
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
Db 205 CACGCCCGVGGCTCTGGTCAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGCATC 261
QY 75 AsnThrCysThrGlyGluValHisGluValArgGluSerPheLeuAlaArgAlaAsn 93
Db 262 AACACCCCGCACCGSGAGTTCACCAACGCAAGGAAAYTCTCTCCAGGCCGAT 318

RESULT 10
US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-23
Alignment Scores:
Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-2 (1-93) x US-10-090-035-23 (1-584)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 46 ATGGCGCAGTTCAGGAGGTGGACTACTCTCGAGGAGTTCAGGCGGTGGCTACCG 105
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
Db 106 GCC-----CGCGCGCGTTCGCGCGGTTCGAGGAGACATCTCAAGGAGAGCGTTC 156

QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHis 57
Db 157 GTGCAGGAGTTCGACACCGCGCGCGC-----CGCCAYGGTCCACCGGTTCAC 204
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
Db 205 CACGCCCGVGGCTCTGGTCAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGCATC 261
QY 75 AsnThrCysThrGlyGluValHisGluValArgGluSerPheLeuAlaArgAlaAsn 93
Db 262 AACACCCCGCACCGSGAGTTCACCAACGCAAGGAAAYTCTCTCCAGGCCGAT 318

RESULT 11
US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc_feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-13
Alignment Scores:
Pred. No.: 1.39e-24 Length: 348
Score: 251.00 Matches: 51
Percent Similarity: 70.59% Conservative: 9
Best Local Similarity: 60.00% Mismatches: 21
Query Match: 49.22% Indels: 4
DB: 9 Gaps: 2

US-10-090-035-2 (1-93) x US-10-090-035-13 (1-348)

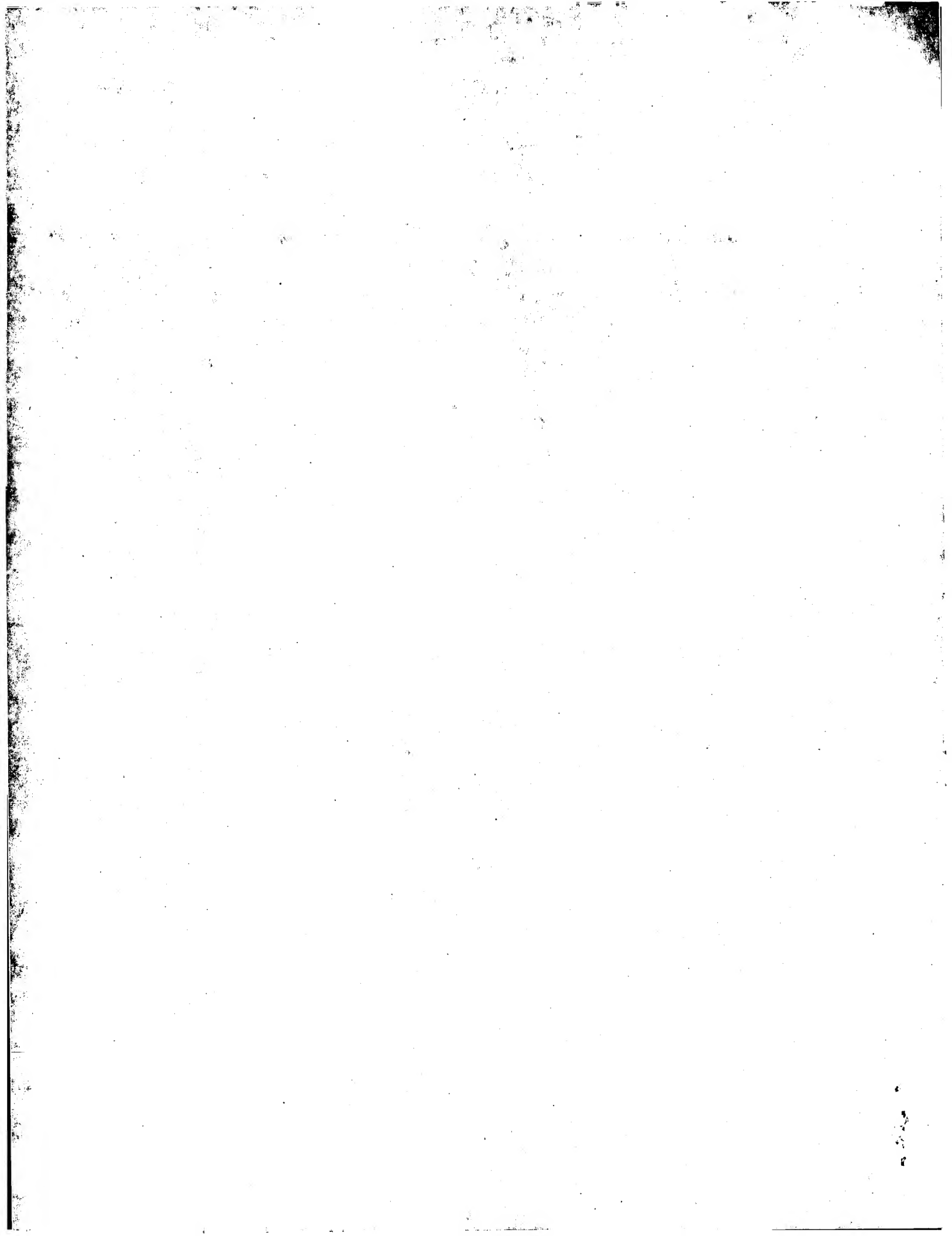
QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db 52 ATGGCTCAGTTCAGGAGGTGGACTACTCTCGAGGAGTTCGCTCGGTACCCGCC 111
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db 112 GCGGGTCTCTCGCGCGCGCGCGCGTTCGAGGAGTTCGCTCGGTACCCGCC 171
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db 172 GAGATCGACANG-----TCCGGCTCCGCGCGCGCGCGCAGCAACACACCCGCC 222
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 223 AACGACTACCTNATGGTTCGCGAGACCAAGGTGAGGAGGACTTTTAAACACCTGCACCGC 282

QY 80 GluValHisGluArg 84
Db 283 GAGTTTCGCGAGCGC 297

RESULT 12
US-09-923-076-2788
; Sequence 2788, Application US/09923876

[illegible]

Search completed: June 16, 2003, 13:51:16
Job time : 132 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:48:33 ; Search time 25 Seconds
(without alignments)
357.620 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYOVEDYCSSEERSVAPA.....INTCTGVEHRESFLARAN 93

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92.5	18.1	196	2 G85435	TINY-like protein
2	92.5	18.1	259	2 T52619	TINY-like protein
3	92	18.0	414	2 A48273	delta/YVI/NF-EI/UC
4	89	17.5	1245	2 T49815	related to multifu
5	86.5	17.0	471	2 T33997	hypothetical prote
6	84	16.5	473	2 D96391	hypothetical prote
7	82.5	16.2	735	2 T45059	hypothetical prote
8	82	16.1	173	2 T51469	glycine/proline-ri
9	81.5	16.0	495	1 S31223	transcription fact
10	81	15.9	86	2 T16437	hypothetical prote
11	81	15.9	191	2 A42844	abscisic acid-and
12	80	15.7	133	2 T09608	environmental stre
13	80	15.7	1273	2 T00338	hypothetical prote
14	79	15.5	204	2 T09592	protein corA, cold
15	79	15.5	414	2 A40350	transcription repr
16	77	15.1	754	2 JC4898	Down-syndrome-crit
17	76.5	15.0	77	2 T16436	hypothetical prote
18	76.5	15.0	332	2 C83682	hypothetical prote
19	76	14.9	102	2 T30119	hypothetical prote
20	75	14.9	1500	2 T00080	transcription fact
21	75.5	14.8	425	1 JH0710	transforming prote
22	75	14.7	369	1 TVEVAF	probable limonene
23	75	14.7	600	2 F71434	dve protein - frul
24	75	14.7	1019	2 T00117	CRAG protein - fru
25	75	14.7	1441	2 T13717	hypothetical prote
26	74.5	14.6	410	2 T26757	hypothetical prote
27	74	14.5	83	2 T16435	hypothetical prote
28	74	14.5	895	2 JC7089	zinc finger bindin
29	74	14.5	1891	2 T13594	hypothetical prote

ALIGNMENTS

RESULT 1

G85435
TINY-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85435
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <STO>
A:Cross-references: GB:NC_001268; NID:g7270639; PIDN:CAB80356.1; GSPDB:GN00140
A:Gene: A74g36900
A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.012;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY 21 GFGHGGG-VQOHVKEKFE-----VDTVSRAGA---NHHHHHGH-HGGHGFVVRER--R 69
Db 99 GGGVGGGDMSAAYTRKAAEVGAQVDALEAAGAGGNRRHHHHHHQHGRNHDYVDNHSYR 158
QY 70 VEEDINTCTGE 80
Db 159 INDDLMECSK 169

RESULT 2

T52619
TINY-like protein [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C:Accession: T52619
R:ferryn, N.; Heijnen, L.; De Keyser, A.; Van Asseldonck, M.; De Clercq, R.; Verbakel
ueller, C.; Mayer, K.; Dehais, P.; Rombauts, S.; Van Montagu, M.; Rouze, P.; Vos, P.
FEBS Lett. 445, 237-245, 1999
A:Title: Evidence for an ancient chromosomal duplication in Arabidopsis thaliana by
A:Reference number: Z26032; MUID:99192287; PMID:10094464
A:Accession: T52619
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-259 <TER>
A:Cross-references: EMBL:AJ002598; PIDN:CAA05630.1
C:Genetics:
A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 259;

Best Local Similarity 33.8%; Pred. No. 0.016; Mismatches 12; Indels 11; Gaps 5; Matches 24; Conservative 12

QY 21 GFGHGGG-VQOHVVKFEE----VDTVSRAGA---NHHHHHGH-HGHHGFFVVRVRET--R 69
Db 162 GGGVNGGDMGSAAYTRRAAEVGAQVDALAEAGAGNRHHHHHHQHQRGHHHDYVDNHSYR 221
QY 70 VEEDINTCTGE 80
Db 222 INDDLMECSSK 232

RESULT 3
A48273
delta/VYI/NF-El/UCRBP transcription factor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C:Accession: A48273; A42055; A56418
R:Safrany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A:Title: Characterization of the mouse gene that encodes the delta/VYI/NF-El/UCRBP trans
A:Reference number: A48273; MUID:93296177; PMID:8516301
A:Accession: A48273
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-414 <RES>
A:Cross-references: GB:113968; NID:9293847; PIDN:AAA40477.1; PID:9293849
R:Flanagan, J.R.; Becker, K.G.; Ennlist, D.L.; Gleason, S.L.; Driggers, P.H.; Levi, B.Z.;
Mol. Cell. Biol. 12, 38-44, 1992
A:Title: Cloning of a negative transcription factor that binds to the upstream conserved
A:Reference number: A42055; MUID:92107191; PMID:1309593
A:Accession: A42055
A:Molecule type: mRNA
A:Residues: 1-414 <FLA>
A:Cross-references: GB:M73963; NID:g202270; PIDN:AAA40522.1; PID:g202271
A:Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:P:74642)
R:Haribaran, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A:Title: delta, a transcription factor that binds to downstream elements in several poly
A:Reference number: A56418; MUID:92052178; PMID:1946404
A:Accession: A56418
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218, 'S', 220-374, 'G', 376-414 <HAR>
A:Cross-references: GB:M74590; NID:g192940; PIDN:AAA37521.1; PID:g192941
C:Genetics:
A:Introns: 227/1; 281/2; 301/3; 354/3
C:Keywords: transcription factor; zinc finger

Query Match 18.0%; Score 92; DB 2; Length 414;
Best Local Similarity 31.0%; Pred. No. 0.031;
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps 2;

QY 21 GFGHGGVQOHVVKFEEVDTVSRAGANHHHHHHGGHGGHGFVVRVREEDINTCTGE 80
Db 55 GGGHGGGGGGGGH-----GHAGHHHHHHHHHHPPMIALQPLVTDD----PTQ 98
QY 81 VHERRESFLAR 91
Db 99 VHHQEVILVQ 109

RESULT 4
T49815
related to multifunctional cyclin-dependent kinase PH085 [imported] - Neurospora crassa
N:Alternate names: protein B24H17.10
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: T49815
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49815

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1245 <SCH>
A:Cross-references: EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.10
A:Experimental source: BAC clone B24H17; strain OR74A
C:Genetics:
A:Gene: NCSP:B24H17.10
A:Map position: 6
A:Introns: 6/2; 1141/3

Query Match 17.5%; Score 89; DB 2; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.21;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 45 SRAGANHHHHHHHHGGHGG 62
Db 901 SRSGTHHHHHHHHHGGHGG 918

RESULT 5
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33997
R:Pauley, A.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: EMBL:AF125964; PIDN:AAD14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 17.0%; Score 86.5; DB 2; Length 471;
Best Local Similarity 34.8%; Pred. No. 0.14;
Matches 24; Conservative 4; Mismatches 16; Indels 25; Gaps 4;

QY 15 RSVAPAGFRHGG-----GVQOHVVKFEEVDTVSRAGANHHHHHHGGHGG 56
Db 372 RSRSPRGG-HGRRHGGPHCPHCRHGGHGGPHH-----HHHDGRSPSRHGGHHHHHHGGCRPF 426
QY 57 --HHGGHGF 63
Db 427 PPHGGHHHF 435

RESULT 6
D96591
hypothetical protein T24C10.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96591
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maizl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <STO>

A:Cross-references: GB:AE005173; NID:99857523; PIDN:AAG00878.1; GSPDB:GN00141
C:Genetics:
A:Gene: T24C10.10
A:Map position: 1

Query Match 16.5%; Score 84; DB 2; Length 473;
Best Local Similarity 36.8%; Pred. No. 0.26; Mismatches 18; Indels 8; Gaps 3;
Matches 21; Conservative 10;

QY 8 DYCEEVRSVAPAGFGHGGVQOHHVYKKEFEVDVTSRAGANHHHHGHG---HGCH 61
DB 409 DYISEFV-SLLPKSIRR---VAEPIPEVQVLEAKAGDDHDDHHGHGHAHACY 460

RESULT 7
T45059
hypothetical protein Y39B68.gg [Imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
R:Wilson, R.; Alnsough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.
Nature 368, 32-38, 1994
A:Authors: Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stock, L.; Wilkinson-Sproat, J.; Wooldman, P.
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID:94150718; PMID:7906398
A:Accession: T45059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-735 <WIL>
A:Cross-references: EMBL:AL128896; NID:96434440; PIDN:CAB60938.1; PID:96434473
A:Experimental source: clone Y39B68
C:Genetics:
A:Map position: 3
A:Introns: 18/1: 69/1
A:Note: Y39B68.gg

Query Match 16.2%; Score 82.5; DB 2; Length 735;
Best Local Similarity 26.4%; Pred. No. 0.6; Mismatches 19; Conservative 1; Indels 43; Gaps 2;
Matches 19;

QY 23 GRHGGVQOHHVYKKEFEVDVTSRAGANHHH-----HGHHGHHHAPHHEHGHGHHGHHGHTHH 662
DB 614 GHGAGYGNH-----HGHHGHHHAPHHEHGHGHHGHHGHTHH 662
QY 54 ---HHGHHGGH 62
DB 663 SLAHRHGGHGG 674

RESULT 8
T51469
glycine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein K10A8_130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
R:Sato, S.; Nakamura, Y.; Kaneo, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <SAT>
A:Cross-references: EMBL:AL391151
A:Experimental source: cultivar Columbia; BAC clone K10A8
C:Genetics:
A:Map position: 5
A:Introns: 97/1
A:Note: K10A8_130

Query Match 16.1%; Score 82; DB 2; Length 173;
Best Local Similarity 34.7%; Pred. No. 0.15; Mismatches 17; Conservative 7; Indels 10; Gaps 2;
Matches 17;

QY 19 PAGFGRHGGVQOHHVYKKEFEVDVTSRAGA---NHHHHHHHHGGHGF 63
DB 93 PSHSGHHGGIGAIAG-----GVAAAGAHMHHHHGHHGHHGHGY 135

RESULT 9
S31223
transcription factor Brn-1 - mouse
N:Alternate names: class III POU domain protein brain-1
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
C:Accession: S31223
R:Hara, Y.; Rovescalli, A.C.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A:Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A:Reference number: S31223; MUID:92228768; PMID:1565620
A:Accession: S31223
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAR>
A:Cross-references: EMBL:M89299; NID:920044; PIDN:AAA39960.1; PID:9200445
C:Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:27-49/Region: glycine-rich
F:101-112/Region: alanine-rich
F:162-180/Region: histidine/proline-rich
F:186-201/Region: alanine-rich
F:236-247/Region: glycine-rich
F:267-291/Region: histidine/proline-rich
F:316-383/Domain: POU domain homology <POU>
F:402-458/Domain: homeobox homology <HOX>

Query Match 16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.5; Mismatches 24; Conservative 7; Indels 25; Gaps 2;
Matches 24;

QY 19 PAGFGRHGGVQOHHVYKKEFEVDVTSRAGANHHHH-----GHHGGHG- 62
DB 237 PGGGGGAGGAGSLVHPLVGRDTPELAEHHHHHHHHHAPHPPHHAQGPPIHGGGA 296
QY 63 -----FVRETREVEDINTCTGEVHERR 85
DB 297 GPGLNSHDSDEDTPTSDLEQFAKQKORR 328

RESULT 10
T16437
hypothetical protein F53A9.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16437
R:Miller, N.
submitted to the EMBL Data Library, March 1995
A:Description: The sequence of C. elegans cosmid F53A9.
A:Reference number: Z19513
A:Accession: T16437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-86 <ML>
A:Cross-references: EMBL:U23523; NID:9746551; PID:9746557; PIDN:AAC46561.1; CBSP:F53A
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CBSP:F53A9.6

Query Match 15.9%; Score 81; DB 2; Length 86;
Best Local Similarity 29.5%; Pred. No. 0.089; Mismatches 23; Conservative 5; Indels 24; Gaps 3;
Matches 23;

QY 1 MAYQEVDCSEVRSVAPAGFGHGG-----GVQHHVYKKEFEVDVTS 45

C:Species: homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
R:Ragase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, N. DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00338
A:Accession:nc001090; translated from GB/EMBL/DBJ

A; Status: preliminary, Translated from: mRNA
A; Molecule type: mRNA
A; Residues: 1-1273 <NAG>
A; Cross-references: EMBL:AB011142; NID: g3043663; PID: g3043664
A; Experimental source: brain; clone HH2365
C: Genetics
G; Note: KIAA0570
G; Species: human
G; Species-family: human hypothetical protein KIAA0570

RESULT 14

T09592
protein corA, cold- and drought-regulated - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09592
R:Laberge, S.; Castonguay, Y.; Vezina, L.P.
Plant Physiol. 101, 1411-1412, 1993
A:Title: New cold- and drought-regulated gene from Medicago sativa.
A:Reference number: Z16754; MUID:94143496; PMID:8310076
A:Accession: T09592
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-204 <LAB>
A:Cross-references: EMBL:L03708; NID:g289122; PID:AAA98933.1; PID:g289123
C:Genetics:
A:Gene: corA

RESULT 12

T09608
environmental stress-induced protein - alfalfa (fragment)
C:Species: Medicago sativa (alfalfa)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Jun-2000
C/Accession: T09608
R:Luo, M.; Mohapatra, S.; Hill, R.
submitted to the EMBL Data Library, February 1992
A:Description: Nucleotide sequence analysis of a environmental stress and ABA inducible
A:Reference number: Z16770
A/Accession: T09608
A>Status: Preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-133 <L00>
A:Cross-references: EMBL:M74190; NID:gl166373; PID:gl166374
A:Experimental source: cultivated Anik
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match	15.7%;	Score 80;	DB 2;	Length 133;
Best Local Similarity	28.7%;	Pred. No. 0.18;		
Matches	27;	Conservative	8;	Mismatches 35; Indels 24; Gaps 4;
QY	21	GFGRGGGQQQHV-VKEKFEEDVTY-----SRAGANH-----HHHHGHH	58	
		:		
Db	35	GHGGGGAESVAVQTEKTNEDYNDKAYGGGSYNDGRGGYNHGGGYNHGGGGHGGH	94	
QY	59	GGHGFVWRE--TRVEEDINTCTGEVHERRESFLA	90	
		: : :		
Db	95	GGHGGTDTVTLTKLRTTLTDNDIIMHHAFLS	128	

A;Molecule type: mRNA
A;Residues: 1-414 <SHI>
A;Accession: CR:W77698. NID:q186767. PIDN:AAA59467.1; PID:q1867678

A; Cross-References: GB.M/7020, 112.3-00, 017

A:Experimental source: HeLa cells
A:Note: the authors translated the codon CGC for residue 371 as Lys, CGA for residue 375
R:Whitson, R.H.; Huang, T.; Dang, J.; Itakura, K.
submitted to the EMBL Data Library, July 1992
A:Description: Observed and predicted DNA binding of a zinc finger protein which recogni
A:Reference number: S78494
A:Accession: S78494
A:Molecule type: mRNA
A:Residues: 1-195,'G',197-414 <WHI>
A:Cross-references: EMBL:Z14077; NID:g38010; PIDN:CAA78455.1; PID:g38011
R:Park, K.; Atchison, M.L.
Proc. Natl. Acad. Sci. U.S.A. 88, 9804-9808, 1991
A:Title: Isolation of a candidate repressor/activator, NF-E1 (YY-1), delta, that binds t
A:Reference number: A56419; MUID:92052179; PMID:1946405
A:Accession: S33712
A:Molecule type: mRNA
A:Residues: 1-64,'R',66-195,'G',197-414 <PAR>
A:Cross-references: GB:M76341; NID:g189173; PIDN:AAA5926.1; PID:g189174
C:Keywords: DNA binding; transcription regulation; zinc finger
F:298-320/Region: zinc finger CCHH motif
F:327-347/Region: zinc finger CCHH motif
F:355-377/Region: zinc finger CCHH motif
F:385-407/Region: zinc finger CCHH motif

Query Match 15.5%; Score 79; DB 2; Length 414;
Best Local Similarity 29.2%; Pred. No. 0.78;
Matches 21; Conservative 6; Mismatches 25; Indels 20; Gaps 3;
Qy 21 GFGRRGGVQOHVVKKEFEVDVSRAGANHHHHHHGGH-GFVRETRVEEDINTCTG 79
Db ||||||| | ||||| | : | :
55 GGGDHGGG-----GGHGHAGHHHHHHHHHHPPMIALQLPLVTD-----PT 95
Qy 80 EVHERRESFLAR 91
Db :|| :| :| :
96 QVHHHQEVILVQ 107

Search completed: June 6, 2003, 12:53:24
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:42:32 ; Search time 13 Seconds
(without alignments)
296.715 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYQEVDCSEVRVAP.....INTCTGVEHRESEFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	18.0	414	1	TYYL_MOUSE
2	81.5	16.0	495	1	BRN1_MOUSE
3	81.5	16.0	497	1	BRN1_RAT
4	80.5	15.8	500	1	BRN1_HUMAN
5	79	15.5	204	1	CORA_MEDSA
6	79	15.5	414	1	TYYL_HUMAN
7	77	15.1	763	1	DYRA_HUMAN
8	77	15.1	763	1	DYRA_MOUSE
9	77	15.1	763	1	DYRA_RAT
10	76	14.9	403	1	MAF_HUMAN
11	76	14.9	551	1	CBX4_MOUSE
12	75.5	14.8	425	1	POU1_BRARE
13	75.5	14.8	558	1	CBX4_HUMAN
14	75	14.7	369	1	TMAF_AVIS4
15	74	14.5	449	1	CSUP_DROME
16	74	14.5	895	1	Z281_HUMAN
17	74	14.5	2190	1	CCAD_CHICK
18	73.5	14.4	147	1	YV59_CAEEL
19	73.5	14.4	684	1	EP84_HCMVA
20	73	14.3	606	1	GAB_DROME
21	73	14.3	633	1	SNF1_YEAST
22	72	14.1	354	1	OTX1_HUMAN
23	71.5	14.0	459	1	A2AC_DIDMA
24	71.5	14.0	532	1	ZIC2_HUMAN
25	71.5	14.0	699	1	SRCH_HUMAN
26	71	13.9	2038	1	FSH_DROME
27	70.5	13.8	139	1	SALA_DROSI
28	70.5	13.8	355	1	OTX1_MOUSE
29	70.5	13.8	355	1	OTX1_RAT
30	70.5	13.8	727	1	CCRL_HORSE
31	70.5	13.8	2175	1	HNCU_DROME
32	70	13.7	190	1	HUNE_DROCR
33	70	13.7	370	1	MAF_MOUSE

RESULT 1

ID	TYYL_MOUSE	STANDARD	PRT	414 AA
AC	Q00899;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Transcriptional repressor protein YY1 (Yin and yang 1) (YY-1) (Delta transcription factor) (NF-E1) (UCR-motif DNA-binding protein).			
GN	YY1 OR UCRBP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92107191; PubMed=1309593;			
RA	Flanagan J.R., Becker K.G., Ennist D.L., Gleason S.L., Driggers P.H.,			
RA	Levi B.-Z., Appella E., Ozato K.;			
RT	"Cloning of a negative transcription factor that binds to the			
RT	upstream conserved region of Moloney murine leukemia virus.;"			
RL	Mol. Cell. Biol. 12:38-44(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=Liver;			
RA	MEDLINE=93296177; PubMed=8516301;			
RA	Safrany G., Perry R.P.;			
RT	"Characterization of the mouse gene that encodes the delta/YY1/NF-			
RT	El/UCRBP transcription factor.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:5559-5563(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92052178; PubMed=1946404;			
RA	Hariharan N., Kelley D.E., Perry R.P.;			
RT	"Delta, a transcription factor that binds to downstream elements in			
RT	several polymerase II promoters, is a functionally versatile zinc			
RL	finger protein.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9799-9803(1991).			
CC	-!- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS			
CC	POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND			
CC	VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION			
CC	START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND			
CC	DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A			
CC	REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. BINDS TO			
CC	THE UPSTREAM CONSERVED REGION (UCR) (5'- GCACATTT-3') OF MOLONEY			
CC	MURINE LEUKEMIA VIRUS (MULV).			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.			
CC	-!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-			
CC	FINGER PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-aib.ch/announce/			

34	69.5	13.6	515	1	KE4L_CAEEL	Q9xtq7 caenorhabdi
35	69.5	13.6	623	1	PNTL_DROME	P51022 drosophila
36	69.5	13.6	726	1	CCTL_HUMAN	O60563 homo sapien
37	69	13.5	85	1	ANTF_SARPE	Q08617 sarcophaga
38	69	13.5	332	1	DLX2_MOUSE	P40764 mus musculus
39	69	13.5	351	1	CAV2_CAEEL	O18879 caenorhabdi
40	69	13.5	359	1	HKN1_MAIZE	P24345 zea mays (m
41	69	13.5	410	1	BR3B_HUMAN	Q12837 homo sapien
42	69	13.5	411	1	BR3B_MOUSE	O63934 mus musculus
43	68.5	13.4	323	1	MAFB_HUMAN	O9y5q3 homo sapien
44	68.5	13.4	873	1	RX_DROME	Q9W2q1 drosophila
45	68	13.3	351	1	HRPX_PLALO	P04929 plasmodium

ALIGNMENTS

```

proc. Natl. Acad. Sci. U.S.A. 69:3260-3264 (1972).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- TISSUE SPECIFICITY: BRAIN.
-!- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CLASS-3 SUBFAMILY
-!- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

CC	EMBL; M88299; AAA39960.1; -	
CC	PIR; S31223; S31223.	
DR	HSSP; P14859; LOCT.	
DR	MGD; MGI:102564; Pou3f3.	
DR	InterPro; IPR001356; Homeobox.	
DR	InterPro; IPR000327; POU_domain.	
DR	Ffam; PF00046; homeobox; 1.	
DR	Ffam; PF00157; pou; 1.	
DR	PRINTS; PR00028; POUDOMAIN.	
DR	ProDom; PD000010; Homeobox; 1.	
DR	ProDom; PD000583; POU_domain; 1.	
DR	SMART; SM00389; HOX; 1.	
DR	SMART; SM00352; POU; 1.	
DR	PROSITE; PS00027; HOMEBOX_1; 1.	
DR	PROSITE; PS00071; HOMEBOX_2; 1.	
DR	PROSITE; PS00035; POU_1; 1.	
DR	PROSITE; PS00465; POU_2; 1.	
KW	Nuclear protein; DNA-binding; Homeobox.	
FT	POLY-GLY.	49
FT	DOMAIN	28
FT	DOMAIN	101
FT	DOMAIN	112
FT	DOMAIN	201
FT	DOMAIN	186
FT	DOMAIN	267
FT	DOMAIN	313
FT	DOMAIN	383
FT	DOMAIN	401
FT	DOMAIN	460
FT	DOMAIN	495
FT	DOMAIN	50012
FT	DOMAIN	50012 MW; 77B802E890C9A014 CRC64;

```

Query Match      16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.26;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;

QY    19  PAGFGRHGGVGQHVYKEKFEEDVTVSRAGANHHHH-----CHHGCG- 62
       ||| :|| | ||| :|||||
Db     237 PGGGGAGGAQLVHPGLVRGDTPELAEHHHHHHHHHAHPHPPHAAQGPPHHGGGA 296

QY    63  -----FWVRETRVEEDINTCTGEVHRR 85
       :|| :|| :||
Db     297 GPGLNSHDPSHDEDTPTSDLEQFAKFQKR 328

```

RESULT 3	BRNI_RAT	STANDARD;	PRT;	497 AA.
ID	BRNI_RAT			
AC	O63262;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1999 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).			
GN	POU3F3 OR BRN1 OR BRN-1 OR RH52.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98070400; PubMed=9405434;			
RA	Schreiber J., Enderich J., Sock E., Schmidt C., Richter-Landsberg C.,			
RA	Wegner M.;			
PT	"Redundancy of class III POU proteins in the oligodendrocyte			

14 **AccountName**

```

or send an email to license@isb-sib.ch).
-----
EMBL; M73963; AAA40522.1; -.
EMBL; L13968; AAA40477.1; -.
EMBL; L13969; AAA40477.1; JOINED.
EMBL; L13965; AAA40477.1; JOINED.
EMBL; L13966; AAA40477.1; JOINED.
EMBL; L13967; AAA40477.1; JOINED.
EMBL; M74590; AAA37521.1; -.
HSP; P25490; IUBD.
DR TRANSFAC; T00278; -.
DR TRANSFAC; T00865; -.
DR MGB; MGI:99150; YL1.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR PRINTS; PR00048; ZINCFINGER.
DR PRODOM; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS01057; ZINC_FINGER_C2H2_2; 4.
DR Transcription regulation; Repressor; Activator; Nuclear protein;
KW transcription binding; DNA-binding; Repeat.
KW zinc-finger; Metal-binding; ASP/GLU-RICH (ACIDIC).
FT DOMAIN 43 53
FT DOMAIN 54 70
FT DOMAIN 71 82
FT DOMAIN 161 170
FT DOMAIN 296 407
FT ZN_FING 296 320
FT ZN_FING 325 347
FT ZN_FING 353 377
FT ZN_FING 383 407
FT DOMAIN 257 341
FT DOMAIN 333 371
FT DOMAIN 371 397
FT CONFLICT 219 219
FT CONFLICT 375 375

```

```
SQ SEQUENCE 414 AA; 44717 MW; C012378288E98AF9 CRC64;  
Query Match      18.0%; Score 92; DB 1; Length 414;  
Best Local Similarity 31.0%; Pred. No. 0.017;  
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps . 2  
  
QY          21 GFGRHGGGVQQHVVKKEFEVDTVSRAGHHHHHHGHHGGFVVRTRVEEDINTCTGE 80  
           ||| | ||||| : : :  
Db          55 GGdHGdGGGGGH-----GHAGHHHHHHHHHHPPMIALQLPLVTDD----PTQ 98  
:  
QY          81 VHERRESFLAR 91  
           || :| :| :
```

Db	99 VHHQEVILVQ 109	RESULT 2	BRNL_MOUSE	STANDARD;	PRT;	495 AA.
ID	BRNL_MOUSE					
AC	P31361;					
DT	01-JUL-1993 (Rel. 26, Created)					
DT	01-JUL-1993 (Rel. 26, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).					
DE	POU3F3 OR OTF8 OR BRN1 OR BRN-1.					
GN	Mus musculus (Mouse)					
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.					
OX	NCBI_Taxid=10090;					
RP	(1)					
RP	SEQUENCE FROM N.A.					
RA	MEDLINE=9228768; PubMed=1565620;					
RA	Hara Y., Rovescalli C., Kim Y., Nirenberg M.;					
RT	"Structure and evolution of four POU domain genes expressed in mouse					
RT	brain.";					

```
RT lineage."
RN J. Biol. Chem. 272:32286-32293(1997).
RL [2]
RP SEQUENCE OF 325-449 FROM N.A.
RC TISSUE-Hypothalamus;
RX MEDLINE-92228769; PubMed-1348859;
RA le Moine C., Young W.S.;
RT "RHS2, a POU domain-containing gene, and its expression in developing
RN and adult rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3285-3289(1992).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM EMBRYONIC DAY 11.5 INTO
CC ADULTHOOD.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-3 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001641; CAA04893.1; -
DR EMBL; M84644; AAA42041.1; -
DR HSSP; P14859; LOCT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00157; pou; 1.
DR PRODOM; PD000028; POU_DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00583; POU_domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 28 49 POLY-GLY.
FT DOMAIN 103 114 POLY-ALA.
FT DOMAIN 135 143 POLY-PRO.
FT DOMAIN 173 178 POLY-PRO.
FT DOMAIN 188 203 POLY-ALA.
FT DOMAIN 238 249 POLY-GLY.
FT DOMAIN 269 280 POLY-HIS.
FT DOMAIN 294 301 POLY-GLY.
FT DOMAIN 315 385 POU.
FT DNA_BIND 403 462 HOMEBOX.
SQ SEQUENCE 497 AA; 50226 MW; 00640505E343ABC2 CRC64;

Query Match
Best Local Similarity 16.08; Score 81.5; DB 1; Length 497;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;

QY 19 PACGFRGGGQVQHVYKEFEVDVTSRAGANHHHH-----GHHGGHG- 62
DB 239 PGGGGGGAGGAGSLSVHPGLVGRDTPDLAEHHHHHHHAHPHPHPHQAQPPHGGGA 298
QY 63 -----FVRETRVEEDINTCTGEVHER 85
DB 299 GPGLSNDHPDSDSDTPTSDLDLEQFAKQKRR 330

RESULT 4
BRN1_HUMAN
ID P20264; P78379; STANDARD; PRT; 500 AA.
DT 01-FEB-1991 (Rel. 17, Created)
```

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).
GN POU3F3 OR BRN1 OR OTF8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96359175; PubMed-8703082;
RA Suniyama K., Washio-Watanabe K., Saitou N., Hayakawa T., Ueda S.;
RT "Class III POU genes: generation of homopolymeric amino acid repeats
RL under GC pressure in mammals.";
RL J. Mol. Evol. 43:170-178(1996).
RN [2]
RP SEQUENCE OF 332-456 FROM N.A.
RC TISSUE-Brain.
RX MEDLINE-89295573; PubMed-2739723;
RA He X., Treacy M.N., Simmons D.M., Ingraham H.A., Swanson L.W.,
RA Rosenfeld M.G.;
RT "Expression of a large family of POU-domain regulatory genes in
RL mammalian brain development.";
RL Nature 340:35-42(1989).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-3 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB001835; BAA19459.1; -
DR PIR; S05042; S05042.
DR HSSP; P14859; LOCT.
DR TRANSFAC; T04469; -
DR Genew; HGNC:9216; POU3F3.
DR MIM; 602480; -
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000327; POU_domain.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00157; pou; 1.
DR PRINTS; PR00028; POU_DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU_domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 28 50 POLY-GLY.
FT DOMAIN 104 115 POLY-ALA.
FT DOMAIN 136 144 POLY-PRO.
FT DOMAIN 174 179 POLY-PRO.
FT DOMAIN 189 204 POLY-ALA.
FT DOMAIN 239 250 POLY-GLY.
FT DOMAIN 270 294 POLY-HIS.
FT DOMAIN 281 304 POLY-GLY.
FT DOMAIN 295 318 POU.
FT DOMAIN 318 388 HOMEBOX.
FT DNA_BIND 406 465 HOMEBOX.
FT CONFLICT 433 433 A -> S (IN REF. 2).
SQ SEQUENCE 500 AA; 50327 MW; E536EFFFAS212319 CRC64;

Query Match
15.8%; Score 80.5; DB 1; Length 500;
```

```

53 HHGHHGGHGGFVVVRETRVE 71
|| ||||| | | :
185 GGRGGHGGHGGAVQTDNTQ 203

DBD

RESULT 6
TYYL_HUMAN STANDARD; PRT; 414 AA.
ID TYYL_HUMAN
P25490; Q14935;
01-MAY-1992 (Rel. 22, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Transcriptional repressor protein YY1 (Yin and yang 1) (YY-1) (Delta
transcription factor) (NF-E1).
DE
TYL
GN Homo sapiens (Human).
OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
OC NCBI_taxid=9606;
RN
[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=92005716; PubMed=1655281;
RX Shi Y., Seto E., Chang L.-S., Shenk T.;
RT "transcriptional repression by YY1, a human GLI-Kruppel-related
protein, and relief of repression by adenovirus E1A protein.";
RN Cell 67:377-388(1991).
[2] SEQUENCE FROM N.A.
RP TISSUE=Foieskin;
RX MEDLINE=92052179; PubMed=1946405;
RN Park K., Atchison M.;
RT "Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta),
that binds to the immunoglobulin kappa 3' enhancer and the
immunoglobulin heavy-chain mu E1 site.";
RN Proc. Natl. Acad. Sci. U.S.A. 88:9804-9808(1991).
[3] SEQUENCE FROM N.A.
RP Whilkson R.H., Huang T., Dang J., Itakura K.;
RN Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
[4] SUBCELLULAR LOCATION.
RP MEDLINE=98152931; PubMed=9493912;
RX McNeil S., Guo B., Stein J.L., Lian J.B., Bushmeyer S., Seto E.,
AT Atchison M.L., Penman S., van Wijnen A.J., Stein G.S.;
RT "Targeting of the YY1 transcription factor to the nucleolus and the
nuclear matrix in situ: the C-terminus is a principal determinant for
nuclear trafficking.";
RN J. Cell. Biochem. 68:500-510(1998).
[5]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 304-414.
RX MEDLINE=97098436; PubMed=8942976;
RN Houbaviy H.B., Usheva A., Shenk T., Burley S.K.;
RT "Cocrystal structure of YY1 bound to the adeno-associated virus P5
initiator";
RN Proc. Natl. Acad. Sci. U.S.A. 93:13577-13582(1996).
[6]
RP STRUCTURE BY NMR OF 353-379.
RX MEDLINE=98308000; PubMed=9642075;
RN Viles J.H., Patel S.U., Mitchell J.B.O., Moody C.M., Justice D.E.,
RA Uppenbrink J., Doyle P.M., Harris C.J., Sadler P.J., Thornton J.M.;
RT "Design, synthesis and structure of a zinc finger with an artificial
beta-turn.";
RN J. Mol. Biol. 279:973-986(1998).
CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND A
DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. FOR
EXAMPLE IT ACTS AS A REPRESSOR IN ABSENCE OF ADENOVIRUS E1A
PROTEIN BUT AS AN ACTIVATOR IN ITS PRESENCE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-

```


DR	EMBL; AF108830; AAD31169.1; -	DE	EC 2.7.1.- (Protein kinase minibrain homolog) (MNBH) (MP86) (Dual specificity YAK1-related kinase).
DR	HSSP; Q00534; 1B18.	DE	DIRKIA OR DYRK.
DR	Genew; HGNC:3091; DYRK1A.	OS	Mus musculus (Mouse).
DR	MIM; 600855; -	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR	DR InterPro; IPR000719; Euk_pkinase.	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DR	DR Pfam; PF00069; pkinase; 1.	OX	NCBI_TaxID=10090;
DR	DR ProDom; PD000001; Euk_pkinase; 1.	RN	[1]
DR	DR SMART; SM00220; S_TKC; 1.	RP	SEQUENCE FROM N.A.
DR	DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	RC	STRAIN-Swiss Webster / NIH;
DR	DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	RX	MEDLINE=97131512; PubMed=9070862;
DR	DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.	RA	Song W.J., Sternberg L.R., Kasten-Sportes C., van Keuren M.L., Lou L.,
DR	DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;	RA	Chung S.H., Slack A.C., Miller D.E., Glover T.W., Chiang P.W.,
KW	ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing;	RA	Kurnit D.W.;
KW	Polymorphism.	RT	*Isolation of human and murine homologues of the Drosophila minibrain
KW	DOMAIN 117 134	RT	gene: human homologue maps to 21q22.2 in the Down syndrome 'critical
FT	DOMAIN 117 134	RT	region'.
FT	DOMAIN 159 479	RL	Genomics 38:331-339(1996).
FT	NP_BIND 165 173	RN	[2]
FT	BINDING 188 188	RN	SUBCELLULAR LOCATION, AND SUBUNIT.
FT	ACT_SITE 287 287	RX	MEDLINE=97224401; PubMed=9070862;
FT	DOMAIN 509 515	RA	Song W.J., Chung S.H., Kurnit D.M.;
FT	DOMAIN 599 602	RT	*The murine Dyk protein maps to chromosome 16, localizes to the
FT	DOMAIN 607 619	RT	nucleus, and can form multimers.
FT	DOMAIN 656 672	RT	Biochem. Biophys. Res. Commun. 231:640-644(1997).
FT	DOMAIN 664 671	RT	*FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
FT	MOD_RES 219 219	CC	NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,
FT	MOD_RES 319 319	CC	THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
FT	MOD_RES 321 321	CC	SUBSTRATES.
FT	VARSPPLIC 70 78	CC	-1- SURCELLULAR LOCATION: Nuclear.
FT	VARSPPLIC 516 529	CC	-1- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
FT	VARSPPLIC 530 763	CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
FT	VARSPPLIC 516 540	CC	MNB/DYRK SUBFAMILY.
FT	VARSPPLIC 541 763	CC	-----
FT	VARSPPLIC 559 584	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	VARSPPLIC 585 763	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	VARSPPLIC 415 415	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	VARSPPLIC 681 681	CC	use by non-profit institutions as long as its content is in no way
FT	VARSPPLIC 32 32	CC	modified and this statement is not removed. Usage by and for commercial
FT	VARSPPLIC 47 47	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
FT	VARSPPLIC 57 57	CC	or send an email to license@sib-sib.ch).
FT	VARSPPLIC 123 123	CC	-----
FT	VARSPPLIC 397 397	CC	EMBL; U58497; AAC52994.1; -
FT	VARSPPLIC 592 592	CC	HSSP; Q00534; 1B18.
FT	VARSPPLIC 679 679	CC	MCD; MGI:1330299; Dyk1a.
FT	VARSPPLIC 763 763	CC	InterPro; IPR000719; Euk_pkinase.
FT	VARSPPLIC 32 32	CC	InterPro; IPR002290; Ser_thr_pkinase.
FT	VARSPPLIC 47 47	CC	Pfam; PF00069; pkinase; 1.
FT	VARSPPLIC 123 123	CC	ProDom; PD000001; Euk_pkinase; 1.
FT	VARSPPLIC 397 397	CC	SMART; SM00220; S_TKC; 1.
FT	VARSPPLIC 592 592	CC	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
FT	VARSPPLIC 679 679	CC	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
FT	VARSPPLIC 763 763	CC	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
FT	VARSPPLIC 32 32	CC	Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
FT	VARSPPLIC 47 47	CC	ATP-binding; Nuclear protein; Phosphorylation.
FT	VARSPPLIC 123 123	CC	ATP-binding; Nuclear protein; BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT	VARSPPLIC 397 397	CC	(POTENTIAL).
FT	VARSPPLIC 592 592	CC	DOMAIN 117 134
FT	VARSPPLIC 679 679	CC	DOMAIN 159 479
FT	VARSPPLIC 763 763	CC	NP_BIND 165 173
FT	VARSPPLIC 32 32	CC	BINDING 188 188
FT	VARSPPLIC 47 47	CC	ACT_SITE 287 287
FT	VARSPPLIC 123 123	CC	DOMAIN 509 515
FT	VARSPPLIC 397 397	CC	DOMAIN 599 602
FT	VARSPPLIC 592 592	CC	DOMAIN 607 619
FT	VARSPPLIC 679 679	CC	DOMAIN 656 672
FT	VARSPPLIC 763 763	CC	DOMAIN 664 671
FT	VARSPPLIC 32 32	CC	MOD_RES 219 219
FT	VARSPPLIC 47 47	CC	MOD_RES 319 319
FT	VARSPPLIC 123 123	CC	MOD_RES 321 321
FT	VARSPPLIC 397 397	CC	SEQUENCE 763 AA; E117DDDC5E8C74F CRC64;
FT	VARSPPLIC 592 592	CC	Query Match 15.1%; Score 77; DB 1; Length 763;
FT	VARSPPLIC 679 679	CC	Best Local Similarity 33.3%; Pred. No. 1.2;
FT	VARSPPLIC 763 763	CC	Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
FT	VARSPPLIC 32 32	CC	-----
FT	VARSPPLIC 47 47	CC	QY 11 SEEVRSVAPAGFRHGG-----GVQOHVKEKEEVDVTSRA-----GANHHHHHG 56
FT	VARSPPLIC 57 57	CC	555 SPQVQRQPAPLWGSQTEAPTQVTVEHPVQETTFHVAPOQNALHHHGHGSSHHHHHHH 614
FT	VARSPPLIC 123 123	CC	QY 57 HGGHGG 62
FT	VARSPPLIC 397 397	CC	615 HHHHGG 620
FT	VARSPPLIC 592 592	CC	-----
FT	VARSPPLIC 679 679	CC	DB 615 HHHHGG 620
FT	VARSPPLIC 763 763	CC	-----
FT	VARSPPLIC 32 32	CC	RESULT 8
FT	VARSPPLIC 47 47	CC	DYRA_MOUSE STANDARD; .PRT; 763 AA.
FT	VARSPPLIC 57 57	CC	ID DYRA_MOUSE
FT	VARSPPLIC 123 123	CC	AC Q61214;
FT	VARSPPLIC 397 397	CC	DT 15-JUL-1999 (Rel. 38, Created)
FT	VARSPPLIC 592 592	CC	DT 15-JUL-1999 (Rel. 38, Last sequence update)
FT	VARSPPLIC 679 679	CC	DT 16-OCT-2001 (Rel. 40, Last annotation update)
FT	VARSPPLIC 763 763	CC	DE Dual-specificity tyrosine-phosphorylation regulated kinase 1A


```
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
QY 11 SEEVSVAPAGFGRHGG-----GVQHVVKKEFEVDVTSRA-----GANHHHHHG 56
DB 555 SPQVRQGFAPLWGSGETAPTQVTVEHPVQETTFHVAPOQNALHHHGHSSHHHHHHH 614
QY 57 HHGGHG 62
DB 615 HHHHHG 620

RESULT 9
DYRA_RAT
ID DYRA_RAT STANDARD; PRT; 763 AA.
AC Q63470;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dual-specificity tyrosine-phosphorylation regulated kinase 1A
DE (EC 2.7.1.1.) (Protein kinase minibrain homolog) (MNBH) (RP86) (Dual
DE specificity YAK1-related kinase).
GN DYRK1A OR DYRK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND MUTAGENESIS OF TYR-219.
RC STRAIN-Sprague-Dawley; TISSUE-Brain;
RX MEDLINE=96216443; PubMed=8631952;
RA Kentrup H., Becker W., Heukelbach J., Wilmes A., Schuermann A.,
RA Huppertz C., Kainulainen H., Joost H.-G.;
RT "Dyrk, a dual specificity protein kinase with unique structural
RT features whose activity is dependent on tyrosine residues between
RT subdomains VII and VIII.";
RL J. Biol. Chem. 271:3488-3495(1996).
RN [2]
REVSIONS.
RA Kentrup H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
CHARACTERIZATION.
RX MEDLINE=98421512; PubMed=9748265;
RA Becker W., Weber Y., Wetzel K., Eilmbter K., Tejedor F.J.,
RA Joost H.-G.;
RT "Sequence characteristics, subcellular localization, and substrate
RT specificity of DYRK-related kinases, a novel family of dual
RT specificity protein kinases.";
RL J. Biol. Chem. 273:25893-25902(1998).
CC -!- FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
CC NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,
CC THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
CC SUBSTRATES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: UBQUITOUS.
CC -!- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MNB/DYRK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79769; CAA56164.1; -
CC HSP; Q00534; IBI8.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
```

```
Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing.
FT DOMAIN 117 134 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT (POTENTIAL).
FT DOMAIN 159 479 PROTEIN KINASE.
FT NP_BIND 165 173 ATP (BY SIMILARITY).
FT BINDING 188 188 ATP (BY SIMILARITY).
FT ACT_SITE 287 287 BY SIMILARITY.
FT DOMAIN 509 515 POLY-SER.
FT DOMAIN 599 602 POLY-HIS.
FT DOMAIN 607 619 POLY-HIS.
FT DOMAIN 656 672 SER/THR-RICH.
FT DOMAIN 664 671 POLY-SER.
FT MOD_RES 219 219 PHOSPHORYLATION (AUTO-).
FT MOD_RES 319 319 PHOSPHORYLATION (AUTO-).
FT MOD_RES 321 321 PHOSPHORYLATION (AUTO-).
FT VARSPIC 70 78 MISSING (IN SHORT ISOFORM).
FT MUTAGEN 219 219 Y->F: REDUCED AUTOPHOSPHORYLATION ON
FT TYROSINE, BUT NO LOSS OF HISTONE
FT PHOSPHORYLATION.
SQ SEQUENCE 763 AA; 85541 MW; CB5EC7EC4C1F9A7 CRC64;
Query Match 15.18; Score 77; DB 1; Length 763;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
QY 11 SEEVSVAPAGFGRHGG-----GVQHVVKKEFEVDVTSRA-----GANHHHHHG 56
DB 555 SPQVRQGFAPLWGSGETAPTQVTVEHPVQETTFHVAPOQNALHHHGHSSHHHHHHH 614
QY 57 HHGGHG 62
DB 615 HHHHHG 620

RESULT 10
MAF_HUMAN
ID_MAF_HUMAN STANDARD; PRT; 403 AA.
AC Q75444; Q9UP93;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor Maf (Proto-oncogene c-maf).
GN MAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98282186; PubMed=9616139;
RA Chesi M., Bersagel P.L., Shonukan O.O., Martelli M.L., Brents L.A.,
RA Chen T., Schrock E., Ried T., Kuehl W.M.;
RT "Frequent dysregulation of the c-maf proto-oncogene at 16q23 by
RT translocation to an Ig locus in multiple myeloma.";
RL Blood 91:4457-4463(1998).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -!- DISEASE: Some forms of multiple myeloma (MM) tumors are
CC characterized by a chromosomal translocation t(14;16)(q32.3;q23)
CC that involves MAF and an Igh locus.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC

```
DR TRANSFAC; T01430;
DR InterPro; IPR004826; TF_Maf.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRU2; 1.
KW Oncogene; Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 163 167 POLY-ALA.
FT DOMAIN 173 182 HIS-RICH.
FT DOMAIN 184 191 POLY-GLY.
FT DOMAIN 214 223 POLY-GLY.
FT DOMAIN 233 240 POLY-GLY.
FT DOMAIN 274 300 BASIC MOTIF.
FT DNA_BIND 302 323 LEUCINE-ZIPPER.
FT DOMAIN 369 AA; 38892 MW; F38GB220ACE50FF6 CRC64;
SQ SEQUENCE 369 AA; 38892 MW; F38GB220ACE50FF6 CRC64;

Query Match 14.7%; Score 75; DB 1; Length 369;
Best Local Similarity 37.8%; Pred. No. 0.93;
Matches 17; Conservative 2; Mismatches 18; Indels 8; Gaps 1;

QY 26 GGGVQOVHVKEFEVDVTSRAGA-----NHHHHHHHGGHG 62
Db 143 GGSVPAEMGSAAVSVAIAAAAGGAPHYHHHHHHHHHGGGG 187

RESULT 15
CSUP DROME STANDARD; PRT; 449 AA.
AC Q9V3A4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catecholamines up protein.
GN Catecholamines up protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Muscota; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99403013; PubMed=10471719;
RA Stathakis D.G., Burton D.Y., McIvor W.E., Krishnakumar S.,
RA Wright T.R., O'Donnell J.M.;
RA "The catecholamines up (Catsup) protein of Drosophila melanogaster
RA functions as a negative regulator of tyrosine hydroxylase activity.";
RA Genetics 153:361-362(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V., Nelson C.R., Miklos G.L.G.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:47:53 ; Search time 50 Seconds
(without alignments)
383.248 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYOEVDCSEVRVAP.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTRMBL_21.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organelle.*
- 10: sp_phage.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_virus.*
- 14: sp_vertebrate.*
- 15: sp_unclassified.*
- 16: sp_virus.*
- 17: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	274	53.7	229	10 Q8S0B1	Q8S0B1 oryza sativ
2	177	34.7	95	10 Q9SM40	Q9SM40 sporobolus
3	141	27.6	102	10 Q8S0B2	Q8S0B2 oryza sativ
4	92.5	18.1	196	10 Q93ZA6	Q93ZA6 arabidopsis
5	92.5	18.1	196	10 Q9SM63	Q9SM63 arabidopsis
6	92.5	18.1	259	10 Q24643	Q24643 arabidopsis
7	89	17.5	1245	3 Q9P543	Q9P543 neurospora
8	86.5	17.0	408	10 Q8S0B4	Q8S0B4 oryza sativ
9	86.5	17.0	471	5 Q9UAY0	Q9UAY0 caenorhabd
10	86	16.9	793	4 Q9Y2L9	Q9Y2L9 homo sapien
11	84	16.5	441	10 Q944Q7	Q944Q7 arabidopsis
12	84	16.5	473	10 Q9FZ33	Q9FZ33 arabidopsis
13	84	16.5	1561	5 Q9W3D2	Q9W3D2 drosophila
14	83	16.3	168	5 Q9VWM5	Q9VWM5 drosophila
15	82.5	16.2	381	13 Q90WV0	Q90WV0 petromyzon
16	82.5	16.2	735	5 Q9NES7	Q9NES7 caenorhabd

17	82	15.1	173	10 Q9LPS9	Q9LPS9 arabidopsis
18	81	15.9	86	5 Q20691	Q20691 caenorhabd
19	81	15.9	109	11 Q9D6B9	Q9D6B9 mus musculu
20	81	15.9	191	10 Q42448	Q42448 medicago sa
21	81	15.9	397	5 Q9NC84	Q9NC84 strongyloce
22	80.5	15.8	347	5 Q9VX61	Q9VX61 drosophila
23	80.5	15.8	396	10 Q9FH24	Q9FH24 arabidopsis
24	80	15.7	133	10 Q9ZRI2	Q9ZRI2 medicago sa
25	80	15.7	1273	4 Q8Q316	Q8Q316 homo sapien
26	79	15.5	612	10 Q94LS9	Q94LS9 oryza sativ
27	79	15.5	612	10 Q8RUM5	Q8RUM5 oryza sativ
28	78	15.3	156	10 Q39754	Q39754 fagus sylv
29	78	15.3	585	5 Q9U4F0	Q9U4F0 drosophila
30	78	15.3	586	5 Q8T8P8	Q8T8P8 drosophila
31	78	15.3	990	5 Q9U4F1	Q9U4F1 drosophila
32	78	15.3	1020	5 Q9U4K1	Q9U4K1 drosophila
33	77.5	15.2	489	17 Q8THX7	Q8THX7 methanosarc
34	77.5	15.2	537	5 Q9VTJ6	Q9VTJ6 drosophila
35	77.5	15.2	566	5 Q9VTJ7	Q9VTJ7 drosophila
36	77.5	15.2	709	3 Q96W52	Q96W52 emeritella
37	77.5	15.2	1693	5 Q9VXV8	Q9VXV8 drosophila
38	77	15.1	354	5 Q8SSW8	Q8SSW8 dictyosteli
39	77	15.1	509	5 Q94888	Q94888 drosophila
40	77	15.1	509	5 Q9VZN8	Q9VZN8 drosophila
41	77	15.1	522	5 Q9XTK9	Q9XTK9 drosophila
42	77	15.1	834	6 Q9GKQ7	Q9GKQ7 ovis aries
43	76.5	15.0	77	5 Q20690	Q20690 caenorhabd
44	76.5	15.0	332	16 Q9KG54	Q9KG54 bacillus ha
45	76.5	15.0	1283	5 Q95RH4	Q95RH4 drosophila

ALIGNMENTS

RESULT 1

Q8S0B1 PRELIMINARY; PRT; 229 AA.
AC Q8S0B1;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE P0470A12.5 protein.
GN P0470A12.5

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003436; BAB90280.1; --
SQ SEQUENCE 229 AA; 25296 MW; D6ED4AA65FFEF61E CRC64;

Query Match 53.7%; Score 274; DB 10; Length 229;
Best Local Similarity 58.1%; Pred. No. 4.9e-23;
Matches 54; Conservative 14; Mismatches 21; Indels 4; Gaps 2;

QY 2 AYQEVDCSEVRVAP-AGFGRHGGVQVHVVKEFEEVDVVSAGANHHHHHGG 60
| :|||||:|||||:| || |||||:|:| :|||:|:
140 AQEKEVDKSEVRVTPGTGLGRGQVQVHVVKETFBQID---RSGRRHHHHNHN 196
:|||||:|||||:|||||:|:| :|||:|:

QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
:|||||:|||||:|||||:|:| :|||:|:

DB 197 DYLMVRETKVEEDNTCTGEFRERKQSFLLKSD 229
:|||||:|||||:|||||:|:| :|||:|:

RESULT 2
Q9SM40

```

ID Q9SM40 PRELIMINARY; PRT; 95 AA.
AC Q9SM40;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Putative glycine-rich protein.
OS Sporobolus stapfianus (Resurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Eragrostidae; Sporobolus.
OX NCBI_TaxID=56623;
RN [1]
RP SEQUENCE FROM N.A.
RA Neale A.D., Blomstedt C.K., Bronson P., Le T.N., Guthridge K.,
RA Evans J., Gaff D.F., Hamill J.D.;
RA "The isolation of lowly-transcribed genes which are induced during
RT desiccation of the resurrection grass Sporobolus stapfianus";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ242802; CAB61838.1; D756DCE2B68DD85B CRC64;
SQ SEQUENCE 95 AA; 10668 MW; 10668 MW; D756DCE2B68DD85B CRC64;

Query Match 34.7%; Score 177; DB 10; Length 95;
Best Local Similarity 57.7%; Pred. No. 1.5e-12;
Matches 45; Conservative 8; Mismatches 15; Indels 10; Gaps 6;

QY 1 MAYTOE-VDYCSEVRVAPAGFRH-GGGVQOHVVKKEFEVDTVSRAGANHHHHHHH 58
DB 1 MAHFKEVMDY--EVTSMKPGFGRGGGGVQGVVVKKEFEVEQVTPGRS-GHGGHR 57

QY 59 G--GCG---FVVRTRVE 71
DB 58 GNGHGGSGHFOARETFE 75

RESULT 3
Q8SOB2 PRELIMINARY; PRT; 102 AA.
AC Q8SOB2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE P0470A12.4 protein.
GN P0470A12.4.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0470A12." to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003436; BAB90279.1; D756DCE2B68DD85B CRC64;
SQ SEQUENCE 102 AA; 11214 MW; 6FF1266B1CDE7768 CRC64;

Query Match 27.6%; Score 141; DB 10; Length 102;
Best Local Similarity 45.8%; Pred. No. 2e-08;
Matches 27; Conservative 8; Mismatches 18; Indels 6; Gaps 1;

QY 3 YQEVNDYCEVRVAPAGFRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHHHHGGH 61
DB 4 YIEVDHCAEEMRPP-----HAGGEHYAVRSEYEEVDEMARGRHHHGGGGGGH 56

RESULT 4
Q93ZA6 PRELIMINARY; PRT; 196 AA.
ID Q93ZA6;
AC Q93ZA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

101-JUN-2002 (TrEMBLrel. 21, Last annotation update)
AT4G36900/C7A10.460.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
Cheuk R., Chen H., Kim C.J., Koeseema E., Meyers M.C., Banh J.,
Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
"Arabidopsis cDNA clones";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY057683; AAL15314.1;
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR ProDom; PD001423; TF_AP2; 1.
DR SEQUENCE 196 AA; 21392 MW; E423D5570745C9CD CRC64;

Query Match 18.1%; Score 92.5; DB 10; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.013; 24; Indels 11; Gaps 5;
Matches 24; Conservative 12; Mismatches 12; Indels 11; Gaps 5;

QY 21 GFGHGGG-VQOHVVKKEFE---VDTVSRAGA---NHHHHHHH-HGCHGVVRET--R 69
DB 99 GGVNGGGDMSAAVIRKRAAEVGAQVDALEAGAGNRRHHHHHQRHNDYVDNHSYR 158

QY 70 VEEDINTCTGE 80
DB 159 INDDLMECSK 169

RESULT 5
Q9SW63 PRELIMINARY; PRT; 196 AA.
ID Q9SW63;
AC Q9SW63;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE TINY-like protein.
GN C7A10.460 OR AT4G36900.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Mayer K.F.X.,
RA Schueller C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
EU Arabidopsis sequencing project;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99707; CAB16766.1;
DR EMBL; AL161590; CAB80356.1;
DR HSPF; O80337; ZGCC
DR InterPro; IPR001471; TF_ERF
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PRO0367; ETHRSEPELMNT.
DR ProDom; PD001423; TF_AP2; 1.
DR SMART; SM00380; AP2; 1.
```

DE Related to multifunctional cyclin-dependent kinase PH085.
GN B24HI7.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL356815; CAB92623.1; -
DR HSSP; P42773; 1IH8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR004129; GDPD.
DR InterPro; IPR004331; SPX.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF03049; GDPD; 1.
DR Pfam; PF03105; SPX; 1.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat; Cyclin; Kinase; Repeat.
SQ SEQUENCE 1245 AA; 137798 MW; 16BD0F6A04596A9F CRC64;

Query Match 17.5%; Score 89; DB 3; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.29;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 45 SRAGANHHHHHGGCGHG 62
|||:||||| ||| ||
DB 901 SRSGTHHHHHHHHGGHG 918

RESULT 8
Q8SOB4
ID Q8SOB4 PRELIMINARY; PRT; 408 AA.
AC Q8SOB4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P0470A12.2 protein.
GN P0470A12.2
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12.2";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP003436; BAB90277.1; -
SQ SEQUENCE 408 AA; 43466 MW; F7E0297877F6D69C CRC64;

Query Match 17.0%; Score 86.5; DB 10; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.15;
Matches 25; Conservative 10; Mismatches 15; Indels 43; Gaps 5;

QY 12 EEVRSVAPAGFGRH-----GGGQQHWVKEEEDVTSRAGA-----NH----- 52
|||:||||| ||| |||: ||| |||
DB 199 EECDAAGGGYGRHHGAGGAVKQHTYKQHQATEGNGCGGYYNRRQAVAVAGGQ 258
53 HHH-----GHH--GGH 61
DB 259 HHYGGATAAAYGNASNKQHTFAAAAGHHSSGGH 291

SEQUENCE FROM N.A.
TISSUE=BRIN;
MEDLINE=99246063; PubMed=10231032;
Nagase T., Ishikawa K., Suyama H., Kikuno R., Hirose M.,
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XIII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
DNA Res. 6:63-70(1999).
EMBL; AB023233; BAA76860.1; -
InterPro; IPR001715; Calponin-like.
InterPro; IPR002114; HPR_Serp_Site.
InterPro; IPR001611; LRR.
InterPro; IPR003592; LRR_Out.
InterPro; IPR003591; LRR_Typ.
Pfam; PF00307; CH; 1.
Pfam; PF00560; LRR; 6.
PRINTS; PR00019; LEURICHRPT.
SMART; SM00033; CH; 1.
SMART; SM00370; LRR; 3.
SMART; SM00369; LRR_Typ; 1.
PROSITE; PS00021; CH; 1.
PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
NON_TER 1
SEQUENCE 793 AA; 87801 MW; OBE99CC48C3B37C CRC64;
Query Match 16.9%; Score 86; DB 4; Length 793;
Best Local Similarity 32.8%; Pred. No. 0.37;
Matches 22; Conservative 7; Mismatches 28; Indels 10; Gaps 2;
QY 6 EVDYCESEVRSVAPAGFGRHGGVQHVVKKEFEE---VDTVSRAGA-----NHHHHH 55
DB 39 ELPSRGERCDPPAGAGRGGAREKMATPGSPFPVLSVATLPLHPHHHHHHH 98
QY 56 GHGCHG 62
DB 99 QHGGTG 105
RESULT 11
Q94407 PRELIMINARY; PRT; 441 AA.
AC Q944Q7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE At1g54990/F14C21.5
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.D., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF24608; XAL11602.1; -
SQ SEQUENCE 441 AA; 48816 MW; 520163FE0A8DE447 CRC64;
Query Match 16.5%; Score 84; DB 10; Length 441;
Best Local Similarity 36.8%; Pred. No. 0.31;
Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;
QY 8 DYCESEVRSVAPAGFGRHGGVQHVVKKEFEEVDTVSRAGANHHHHHHH---HGHH 61
DB 39 ELPSRGERCDPPAGAGRGGAREKMATPGSPFPVLSVATLPLHPHHHHHHH 98

Db 377 DYISEFV-SLLPKSIRR-----VAEPIPEEVQKVL EEAKAGDDHDDHHGHGHAGY 428

RESULT 12

ID	Q9F233	PRELIMINARY;	PRT;	473 AA.			
DT	Q9F233;						
AC	01-MAR-2001	(T-EMBLrel. 16, Created)					
DC	01-MAR-2001	(T-EMBLrel. 16, Last sequence update)					
DT	01-MAR-2002	(T-EMBLrel. 20, Last annotation update)					
DE	T24C10.10 protein (Hypothetical 52.4 kDa protein).						
GN	T24C10.10 OR F14C21.51 OR At1G54990.						
OS	Arabidopsis thaliana (Mouse-ear cross).						
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;						
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsi.						
NCBI_TaxID=3702;							
[1]	SEQUENCE FROM N.A.						
RP	Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,						
RA	Altai H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,						
RA	Buehler E., Johnson-Hopson C., Chlou J., Choi E., Gonzalez A.,						
RA	Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,						
RA	Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,						
RA	Toriiumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;						
RL	Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.						
[2]	SEQUENCE FROM N.A.						
RP	STRAIN-CV. COLUMBIA;						
RC	MEDLINE-21016719; PubMed-11130712;						
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,						
RA	White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,						
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,						
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,						
RA	Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,						
RA	Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,						
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,						
RA	Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,						
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.-P.,						
RA	Lin X., Liu S.X., Liu Z.A., Luros J.S., Malti R., Marziani A.,						
RA	Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,						
RA	Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,						
RA	Sakano H., Salzerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,						
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,						
RA	Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,						
RA	Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;						
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis						
RT	thaliana.;"						
RT	Nature 408:816-820(2000).						
[3]	SEQUENCE FROM N.A.						
RP	Yamada K., Banph J., Chan M.M., Chang C.H., Chang E., Dale J.M.,						
RA	Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,						
RA	Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,						
RA	Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,						
RA	Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,						
RA	Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,						
RA	Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,						
RA	Davis R.W., Ecker J.R., Theologis A.;						
RT	"Full length cDNA of gene At1g54990 (GI:15221965).";						
RT	Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.						
DR	EMBL; AC064840; AAG00878.1; -						
DR	EMBL; AC069144; AAG51115.1; -						
DR	EMBL; AY074334; AAL67030.1; -						
DR	Hypothetical protein.						
DR	SEQUENCE 473 AA; 52419 MW; D09124A11565BB23 CRC64;						
SQL							

```
Query Match      . 16.5% ; Score 84; DB 10; Length 473;
Best Local Similarity 36.8%; Pred. NO. 0.34;
Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;

      8 DYCSEEVRSVAPAGFGRHGGGVOOHVVREKFEVDVTVSACANHHHHHH-- --HGHH 61
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 409 DYISEFV-SLLPKSIRR---VAEEPIEEVQKVLEEAKAGDDHHHGHGHAGY 460

RESULT 13

Q9W3D2 PRELIMINARY; PRT; 1561 AA.

ID Q9W3D2;

AC Q9W3D2;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE CRAG protein.

GN CRAG OR CG12737.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=107311132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.J., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,

RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Balow R.M., Basu A., Bonkensdale J., Boyaraktiroglou L., Besley E.M.,

RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fessler C., Gabrielian A.E., Garq N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Spier E., Sidlen-Klamis I., Simpson M., Skupski M.P., Smith T.,

RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Xeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

RE EMBL: AE003444; RAP46397.1;

DR FlyBase: FBgn0025864; Crag.

DR InterPro: IPR005112; dDENN.

DR InterPro: IPR001194; DENN.

DR InterPro: IPR005113; uDENN.

DR Pfam: PF03455; dBENN; 1.

DR Pfam: PF02141; DENN; 1.

DR Pfam: PF03456; uDENN; 1.

SQ SEQUENCE 1561 AA; 174616 MW; 10BFDF38A3DD4FC4E CRC64.

Query Match 16.5%; Score 84; DB 5; Length 1561;
Best Local Similarity 23.9%; Pred. No. 1.4;
Matches 28; Conservative 17; Mismatches 30; Indels 42; Gaps 5;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:02:07 ; Search time 1293.5 seconds
(without alignments)
1164.422 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYQVDYDCSEVRVAPV.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUN62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10090035/runat_06062003_105504_10970/app_query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosun62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTENT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10090035@cgn_1_1_2463@runat_06062003_105504_10970 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_htc:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_htc:.*
12: gb_est3:.*
13: gb_est4:.*
14: gb_est5:.*
15: em_estfun:.*
16: em_estom:.*
17: gb_gss:.*
18: em_gss_hum:.*
19: em_gss_inv:.*
20: em_gss_pln:.*
21: em_gss_vrt:.*
22: em_gss_fun:.*
23: em_gss_mam:.*
24: em_gss_mus:.*
25: em_gss_other:.*
26: em_gss_pro:.*
27: em_gss_rod:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	510	100.0	417	13	BM501439	BM501439 PAC0000000
2	505	99.0	465	10	BE129897	BE129897 945032C12
3	505	99.0	500	10	AW787732	AW787732 945002E06
4	505	99.0	523	10	AW331212	AW331212 707049E04
5	505	99.0	524	10	AW288876	AW288876 707009E07
6	499.5	97.9	466	9	A1964534	A1964534 496013D01
7	499.5	97.9	481	9	A1964458	A1964458 496010H04
8	499.5	97.9	539	9	A1855425	A1855425 603016F02
9	499.5	97.9	546	9	AA979839	AA979839 MEST2-B7
10	493.5	97.9	550	12	BG840383	BG840383 MEST12-H1
11	499.5	97.9	553	11	AY104409	AY104409 zea mays
12	499.5	97.9	648	14	BQ619167	BQ619167 RNOSEQ4E0
13	499.5	97.9	648	14	BQ619315	BQ619315 RNOSEQ6C0
14	499.5	97.9	648	14	BQ619318	BQ619318 RNOSEQ6E0
15	499.5	97.9	648	14	BQ619337	BQ619337 RNOSEQ6B1
16	499.5	97.9	648	14	BQ619383	BQ619383 RNOSEQ7B0
17	499.5	97.9	648	14	BQ619390	BQ619390 RNOSEQ780
18	498	97.6	326	10	BE025303	BE025303 945028B09
19	496	97.3	443	10	BE129644	BE129644 945027E06
20	496	97.3	552	10	BE025302	BE025302 945028B09
21	494	96.9	476	10	BE519299	BE519299 945007B05
22	492	96.5	496	10	AW787314	AW787314 945002E06
23	482	94.5	433	10	BE225008	BE225008 945042F02
24	468	91.8	436	10	AW787315	AW787315 945002E06
25	423.5	83.0	399	13	BM318672	BM318672 P11_16_C0
26	423.5	83.0	466	10	BE597738	BE597738 P11_16_C0
27	423.5	83.0	485	10	BE599123	BE599123 P11_16_C0
28	423.5	83.0	513	14	BQ280709	BQ280709 WHE3004_A
29	423.5	83.0	516	10	BE364814	BE364814 P11_16_C0
30	423.5	83.0	519	10	AW680016	AW680016 WSL_34_H1
31	423.5	83.0	522	10	AW679969	AW679969 WSL_34_H1
32	423.5	83.0	537	10	AW745400	AW745400 WSL_34_H1
33	423.5	83.0	541	10	AW745436	AW745436 WSL_34_H1
34	417.5	81.9	348	10	BE593507	BE593507 WSL_100_B
35	410.5	80.5	509	14	BQ280894	BQ280894 WHE3006_B
36	395	77.5	401	10	AW289056	AW289056 707005E07
37	368.5	72.3	274	10	AW288875	AW288875 707000E07
38	367.5	72.1	311	12	BF729420	BF729420 1000077C0
39	308	60.4	453	10	AW923922	AW923922 WSL_30_A1
40	308	60.4	554	10	AW679915	AW679915 WSL_33_G0
41	308	60.4	560	10	AW924079	AW924079 WSL_49_G0
42	308	60.4	566	10	AW677917	AW677917 WSL_12_D0
43	308	60.4	566	10	AW746383	AW746383 WSL_49_G0
44	308	60.4	572	10	AW679713	AW679713 WSL_30_A1
45	308	60.4	572	10	AW925014	AW925014 WSL_74_D0

ALIGNMENTS

RESULT 1
BM501439
LOCUS BM501439
DEFINITION PAC000000000593 Pioneer AF-1 array zea mays linear EST 14-FEB-2002
ACCESSION BM501439
VERSION BM501439.1
KEYWORDS EST.
SOURCE zea mays
ORGANISM zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 417)
AUTHORS Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and Jung,R.

Maize opaque endosperm mutations create extensive changes in
Patterns of gene expression
Unpublished (2002)
Contact: Jung R
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.

FEATURES
source
1. .417
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/notes="Vector: pSport1; Site_1: SalI; Site_2: NotI"

BASE COUNT 81 a 150 c 133 g 53 t
ORIGIN

Alignment Scores:
Pred. No.: 5.6e-47 Length: 417
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-090-035-2 (1-93) x BM501439 (1-417)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 85 ATGGCTTACACAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGGCCCGGCC 144
|||||
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 145 GGCCTCGCGCCGCCACGGCGCGGCTCCAGCAGCAGCTCTCAAGGAGAGTTCGAGGAG 204
|||||
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyGlyGly 60
|||||
Db 205 GTCCAGACGGGTCTCACGGCGCGGCCACACCCACCACCATGTGTACACCGCGGC 264
|||||
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
|||||
Db 265 CACGGCTTCGTGGTGGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGGAG 324
|||||
QY 81 ValHisGluArgArgGluSerPheIleuAlaArgAlaAsn 93
|||||
Db 325 GTCCACGAGCGCAGGAGAGGCTTCTCTCCAGGCGCTAAC 363
|||||

RESULT 2
BE129897 465 bp mRNA linear EST 21-JUN-2000
LOCUS 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897
VERSION BE129897.1 GI:8577260
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 465)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
FEATURES
source
1. .465
/location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 88 a 151 c 149 g 77 t
ORIGIN

Alignment Scores:
Pred. No.: 2.26e-46 Length: 465
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-2 (1-93) x BE129897 (1-465)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 40 ATGGCTTACTACACAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 99
|||||
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 100 GGCCTCGCGCCGCCACGGCGCGGCTCCAGCAGCAGCTCTCAAGGAGAGTTCGAGGAG 159
|||||
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyGlyGly 60
|||||
Db 160 GTCCAGACGGGTCTCACGGCGCGGCCAACCCACCACCATGTGTACACCGCGGC 219
|||||
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
|||||
Db 220 CACGGCTTCGTGGTGGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGGAG 279
|||||
QY 81 ValHisGluArgArgGluSerPheIleuAlaArgAlaAsn 93
|||||
Db 280 TTCCACGAGCGCAGGAGAGGCTTCTCTCCAGGCGCTAAC 318
|||||

RESULT 3
AW787732 500 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787732
VERSION AW787732.1 GI:7844510
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 500)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.

[illegible]

```

1. 500
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mix
same as 707 (SK)"
/tissue_type="cassel,
/dev_stage="fully-gro
/lab_host="pH10a"

```

/note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAL10; size: 1; EColi: CDNA library from tully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
159 c 156 g 88 t

BASE COUNT	97 a	159 c	156 g	for a	for a
given to Library 707				88 t	88 t

Alignment Scores:

Pred. No.:	2,458-46	Length:	500
Score:	505.00	Matches:	92
Percent Similarity:	98.92%	Conservative:	0
Best Local Similarity:	98.92%	Mismatches:	1
Query Match:	99.02%	Indels:	0
DB:	10	Gaps:	0

US-10-090-035-2 (1-93) x AW787732 (1-500)

1	MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla	20
35	ATGCCTACTACCAAGAGGTGGACTACTGTCGAGGAGCGTAGTGCGGCCGGCC	94
21	GlyPheGlyArfHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu	40
95	GGCTTCGGCCACCACGGCGGGCTCCACGACCGCTCGTAAGGAGAAGTTTCGAGGAG	154
41	ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisHisGlyGly	60
155	GTCGACACGGGTCTCACGCGCGGGCGCAACCAACCACCATGGTCACCACGCGGC	214
61	HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu	80
215	CACGGCTTCGTGGTCGAGACCGAGGTTCGAAGAGGACATCAACACCTGCACCGCGAG	274
81	ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn	93
275	TTCACGAGCGCAGGAGAGCTTCCTCGCAGCGCTAAC	313

RESULT 4

	AW331212	523 bp	mRNA	linear	EST 31-JAN-2000
	707049E04.x1	707 - Mixed	adult tissues from Walbot lab (SK)	Zea mays cDNA, mRNA sequence.	
LOCUS	DEFINITION				

AW331212
AW331212.1
EST.
GI:6827569

SOURCE	ORGANISM	Zea mays.	Zea mays	Zea mays
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

REFERENCE
1 (bases 1 to 523)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL	Unpublished (1999)
COMMENT	Contact: Walbot V Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 733 2227
Fax: 650 725 8221
Email: waibotest@stanford.edu
Plate: 707049 row: E column: 04.

	FEATURES	SOURCE
1.	100% Pure Cotton	USA
2.	Machine Washable	India
3.	Soft & Comfortable	Bangladesh
4.	Durable & Long Lasting	Pakistan
5.	Eco-Friendly	Vietnam
6.	Available in Multiple Colors	Thailand
7.	Lightweight & Breathable	Philippines
8.	Easy to Fold & Store	Malaysia
9.	Hypoallergenic	Singapore
10.	Perfect for Travel	Japan

```

1. ...523
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
) )"
/tissue_types="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pCAD10; Site:1: Ecoli; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

```

BASE COUNT	125 a	158 c	149 g	91 t	Unidire
					husk, foot, leaf):

Alignment Scores:

Pred. No.:	2,588-46	Length:	523
Score:	505.00	Matches:	92
Percent Similarity:	98.92%	Conservative:	0
Best Local Similarity:	98.92%	Mismatches:	1
Query Match:	98.02%	Indels:	0
DB:	10	Gaps:	0

US-10-090-035-2 (1-93) x AW331212 (1-523)

1	QY	MetAla	Ty	Tr	Gln	Glu	Val	Asp	Tyr	Cys	Ser	Glu	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20
35	Db	ATGGCTT	ACT	ACC	AGG	AGT	GG	ACT	TCT	TCG	GAG	GAG	TGT	GAG	TCT	CGT	GGC	CCC	GGC	94	
21	QY	Gly	Phe	Gly	Arg	His	Gly	Gly	Val	Gln	His	Val	Val	Lys	Glu	Lys	Phe	Glu	Glu	40	
95	Db	GGCTTC	GGCCG	CAC	GGCGG	CGT	CTC	AGC	AGC	CTG	CTC	AAG	GAG	AAG	ATT	TCG	AGG	AG	154		
41	QY	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	His	His	His	Gly	His	Gly	Gly	60		
155	Db	GTCA	CAC	AGG	TCT	CAC	CGCGG	CGG	CAC	CAC	CAC	CAC	CAC	CAT	GTG	TCA	CCG	CGG	214		
61	QY	His	Gly	Phe	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Leu	Asp	Ile	Asn	Thr	Cys	Thr	80	
215	Db	CAC	GGCT	TCGT	GTG	TCG	CGG	AGAC	CAC	AGG	TCG	AGG	GAC	ATC	AAC	ACC	TGC	ACG	CGG	274	
81	QY	Val	His	Glu	Arg	Arg	Glu	Ser	Phe	Leu	Ala	Arg	Ala	Asn	93						
275	Db	TTCC	CAC	AGG	CGC	AGG	AGAG	CGT	TCCT	CGC	CC	AGG	CGT	TAAC	313						

RESULT 5

RESULT 5
AW288876
LOCUS
AW288876 524 bp mRNA linear EST 16-JAN-2000
DEFINITION
707009E07.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW288876
VERSTON AW300076

KEYWORDS

SOURCE Zea mays.

ORGANISM Zea mays

Eukaryotes
Spermatop

clade; pa

REFERENCE 1 (bases 1-164) W. J. BOYD, W.

AUTHORS	WALDOE, V.
TITLE	Maize EST

University

JOURNAL
COMMENT
Unpublish
Contest

COMMENT:

CONTACT:

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496013 row: D column: 01.

FEATURES
source
1. .466
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.Coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+)"
Wang/Bohnert"

BASE COUNT 76 a 151 c 154 g 85 t

ORIGIN

Alignment Scores:
Pred. No.: 9.2e-46 Length: 466
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI964534 (1-466)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
416 ATGGCTTACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTCGGTGGCCCGGCC 357
Db

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
356 GGCTTCGGCCGCCAGCGCGCGGCGCTCCAGCAGCAGCGCTCGTCAAGGAGATTCGAGGAG 297
Db

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisGly 59
296 GTCGACACGGCTCTACGCGCGCGCGCCAAACACCACCACCATGGTCACCGCGC 237
Db

QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
236 GGCCACGGCTCTGCTGTCGCGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 177
Db

QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
176 GAGTCCACGAGCGCAGGAGAGCTTCTCCGCGAGGCTAAC 135
Db

RESULT 7
AI964458/c 481 bp mRNA linear EST 20-AUG-1999
LOCUS
DEFINITION 496010H04.x1 496 - stressed shoot cDNA library from Wang/Bohnert
lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964458
VERSION AI964458.1 GI:5757171
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 481)
REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES
source
1. .524
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH108"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 107 a 170 c 157 g 89 t

ORIGIN

Alignment Scores:
Pred. No.: 2.58e-46 Length: 524
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-2 (1-93) x AW288876 (1-524)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
73 ATGGCTTACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTCGGTGGCCCGGCC 132
Db

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
133 GGCTTCGGCCGCCAGCGCGCGGCGCTCCAGCAGCAGCGCTCGTCAAGGAGATTCGAGGAG 192
Db

QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisGlyHisGly 60
193 GTCGACACGGCTCTACGCGCGCGCGCCAAACACCACCACCATGGTCACCGCGC 252
Db

QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
253 CACGCGCTCTGCTGTCGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGGAG 312
Db

QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
313 TTCCACGAGCGCAGGAGAGCTTCTCCGCGAGGCTAAC 351
Db

RESULT 6
AI964534/c 466 bp mRNA linear EST 20-AUG-1999
LOCUS
DEFINITION 496013D01.x1 496 - stressed shoot cDNA library from Wang/Bohnert
lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964534
VERSION AI964534.1 GI:5757247
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 466)
REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 723 2221
Email: walbot@stanford.edu
Plate: 496010 row: H column: 04.

FEATURES
source
1. .481
Location/Qualifiers

/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+); XR; Wang/Bohnert"

BASE COUNT 82 a 148 c 154 g 97 t
ORIGIN

Alignment Scores:
Pred. No.: 9,53e-46 Length: 481
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI964458 (1-481)

QY 1 MetalaTyrTrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 451 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGGAGGTGAGTCTGGTGGCCCGGCC 392
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluGlu 40
Db 391 GCCTTCGGCCGCCAGGAGGCGGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTCGAGGAG 332
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59
Db 331 GTCGACACGGTCTCAGCGCGCGCGCCACACACACACACACATGTCACACGGC 272
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 271 GGCACCGGCTTCGTGGTGGCGAGACCGAGGTCGAGGAGGACATCAACACCTGCACGGC 212
QY 80 GluValHisGluArgAlaGluSerPheLeuAlaArgAlaAsn 93
Db 211 GAGGTCCAGGAGCGAGGAGGAGTCTCTCGCCAGGGCTAAC 170

RESULT 8
AI855425/c
LOCUS 603016F02.x1 603 - stressed root cDNA library from Wang/Bohnert lab
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION AI855425
VERSION AI855425.1 GI:5499558
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
COMMENT clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 539)
Walbot, V.

JOURNAL
COMMENT Maize ESTs from various cDNA libraries sequenced at Stanford
University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.

FEATURES
source
1. .539
Location/Qualifiers

/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+); XR; Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 91 a 157 c 172 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 1.08e-45 Length: 539
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI855425 (1-539)

QY 1 MetalaTyrTrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 477 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGGAGGTGAGTCTGGTGGCCCGGCC 418
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluGlu 40
Db 417 GGCTTCGGCCGCCAGGAGGCGGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTCGAGGAG 358
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59
Db 357 GTCGACACGGTCTCAGCGCGCGCGCCACACACACACACACATGTCACACGGC 298
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 297 GGCACCGGCTTCGTGGTGGCGAGACCGAGGTCGAGGAGGACATCAACACCTGCACGGC 238
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 237 GAGGTCCAGGAGCGAGGAGGAGTCTCTCGCCAGGGCTAAC 196

RESULT 9
AA979839
LOCUS MEST2-B7.TW1412.Seq ISUM2 Zea mays cDNA clone linear EST 26-MAY-1998
DEFINITION sequence.
ACCESSION AA979839
VERSION AA979839.1 GI:3157217
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
COMMENT clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 546)
Wen, F.J., Ashlock, D.A. and Schnable, P.S.

JOURNAL
COMMENT Expressed Sequence Tags from B73 Maize Seedlings
Unpublished (1997)
Contact: Schnable, PS
Schnable Laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975
Fax: (515)-294-2299
Email: schnable@iastate.edu

Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206705.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

AUTHORS
TITLE
JOURNAL
COMMENT

PCR Primers
FORWARD: tw1412 (5'-GAAGATACCCACCAACC-3')
BACKWARD: T7-YU (5'-TAATACGACTCATATAGGC-3')
Plate: MEST2 row: B column: 7
Seq primer: tw1412 (5'-GAAGATACCCACCAACC-3').

FEATURES
source
1..546
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST2-B7"
/tissue_type="above ground tissues"
/dev_stage="two-leaf-stage green seedling"
/lab_host="XLI-MFR Blue"
/note="Organ: green seedlings; Vector: pAD-GAL4; Site: 1:
EcoRI; Site: 2: XhoI; ds-cDNA molecules were generated as
follows. First-strand cDNA was prepared from oligo-dT
selected mRNA by priming with an XhoI oligo-dT primer. The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with XhoI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and XhoI sites of the HybriZAP lambda vector
(Stratagene) and excised as pAD-GAL4 phagemids."

BASE COUNT 130 a 168 c 159 g 88 t 1 others
ORIGIN
Alignment Scores:
Pred No.: 1..1e-45 Length: 546
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AA979839 (1-546)
QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 66 ATGGCTTACTACCAAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 125
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 126 GGCTTCGGCCGCCACGAGGGGGGCGTCCAGCAGCACGTCTCAAGAGAGAGTTCGAGGAG 185
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
Db 186 GTCGACACGGTCTCACGCGCGCGGCCAACCCACCACCATGTCACCGCC 245
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 246 GGCACAGGCTTCGTGGTGGCGGACACAGGTCGAGGAGGACATCAACACCTGCACCGGC 305
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 306 GAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGCTAAC 347

RESULT 10
BG840383
LOCUS
DEFINITION
MEST12-H11.T7-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 550)

US-10-090-035-2 (1-93) x BG840383 (1-550)
QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 75 ATGGCTTACTACCAAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 134
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 135 GGCTTCGGCCGCCACGAGGGCGCGTCCAGCAGCACGTCTCAAGAGAGTTCGAGGAG 194
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
Db 195 GTCGACACGGTCTCACGCGCGCGGCCAACCCACCACCATGTCACCGCC 254
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 255 GGCCACGGCTTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 314
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 315 GAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGCTAAC 356

RESULT 11
AY104409

LOCUS AY104409 553 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PC0124784 mRNA sequence.
 ACCESSION AY104409
 VERSION AY104409.1 GI:21207487
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 553)
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
 Arthur, L.W., Hanafey, M., Morgan, M. and Tingey, S.V.
 TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 553)
 AUTHORS Coe, E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 FEATURES
 source
 Location/Qualifiers
 1..553
 /organism="Zea mays"
 /db_xref="taxon:637271"
 /clone_lib="PC0124784"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

BASE COUNT 128 a 173 c 162 g 90 t

Alignment Scores:
 Pred. No.: 111e-45 Length: 553
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 11 Gaps: 1

US-10-090-035-2 (1-93) x AY104409 (1-553)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 76 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGGTGGCGCGGCC 135
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 136 GGCTTCGGCGGCCACGAGGCGCGTCCAGCAGCAGTCTCGAGGAGTTCGAGGAG 195
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
 |||||
 Db 196 GTCGACACGTCTCACCAGCGCGCGGCCAACCCACCACCACCATGTCACACCGC 255
 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspThrCysThrGly 79
 |||||
 Db 256 GGCACACGCTCTGTTGGTGGCGAGACAGGGTCTCGAGGAGGACATCAACACCTGCACCGGC 315
 QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 316 GAGTTCACGAGCGCGAGGAGAGCTTCCTCGCCAGGCGCTAAC 357

RESULT 12

LOCUS BQ619167 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ4E05_SK.abl Salt stressed Zea mays roots cDNA library Zea
 mays cDNA clone RN05EQ4E05_SK.abl similar to No homology, mRNA

sequence.
 ACCESSION BQ619167
 VERSION BQ619167.1 GI:21621161
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Wang, H. and Bohnert, H.J.
 TITLE Genomics of plant stress tolerance
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mark Fredricksen
 Department of Plant Biology
 University of Illinois
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 217/2653473
 Email: bohnertlab@life.uiuc.edu.
 FEATURES
 source
 Location/Qualifiers
 1..648
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone="RN05EQ4E05_SK.abl"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /tissue_type="Roots"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150
 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

Alignment Scores:
 Pred. No.: 1.33e-45 Length: 648
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619167 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 82 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGGTGGCGCGGCC 141
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 142 GGCTTCGGCGGCCACGAGGCGCGTCCAGCAGCAGTCTCGTCAAGGAGAGTTCGAGGAG 201
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
 |||||
 Db 202 GTCGACACGTCTCACCAGCGCGCGGCCAACCCACCACCACCATGTCACACCGC 261
 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspThrCysThrGly 79
 |||||
 Db 262 GGCACACGCTCTGTTGGTGGCGAGACAGGGTCTCGAGGAGGACATCAACACCTGCACCGGC 321
 QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 322 GAGTTCACGAGCGCGAGGAGAGCTTCCTCGCCAGGCGCTAAC 363

RESULT 13

LOCUS BQ619315 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ6C06_SK.abl Salt stressed Zea mays roots cDNA library Zea
 mays cDNA clone RN05EQ6C06_SK.abl similar to No homology, mRNA
 sequence.
 ACCESSION BQ619315
 VERSION BQ619315.1 GI:21621309
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.

Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

Location/Qualifiers
1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/dev_stage="2 weeks old"
/note="vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619318 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 82 ATGGCTTACTACCGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValGlyGluLysPheGluGlu 40
Db 142 GGCCTTCGGCGCCACCGAGGCGGCTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisGly 59
Db 202 GTCCACAGGCTCTACCGCGCGCGCCACACACCACACCATGTCACCGAGG 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGAGGAGAGCTTCCTCGCCAGGCTTAC 363

RESULT 15
BQ619337
LOCUS
DEFINITION
BQ619337 648 bp mRNA linear EST 27-JUN-2002
mays cDNA clone RNOSEQ6E12.SK.ab1 Salt stressed Zea mays roots cDNA library Zea sequence.

ACCESSION BQ619337
VERSION BQ619337.1 GI:21621331
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Wang, H. and Bohnert, H.J.
AUTHORS Genomics of plant stress tolerance
TITLE Unpublished (2002)
JOURNAL Contact: Mark Fredricksen
COMMENT Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

Location/Qualifiers
1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/dev_stage="2 weeks old"
/note="vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619315 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 82 ATGGCTTACTACCGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValGlyGluLysPheGluGlu 40
Db 142 GGCCTTCGGCGCCACCGAGGCGGCTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisGly 59
Db 202 GTCCACAGGCTCTACCGCGCGCGCCACACACCACACCATGTCACCGAGG 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGAGGAGAGCTTCCTCGCCAGGCTTAC 363

RESULT 14
BQ619318
LOCUS
DEFINITION
BQ619318 648 bp mRNA linear EST 27-JUN-2002
mays cDNA clone RNOSEQ6D01.SK.ab1 Salt stressed Zea mays roots cDNA library Zea sequence.

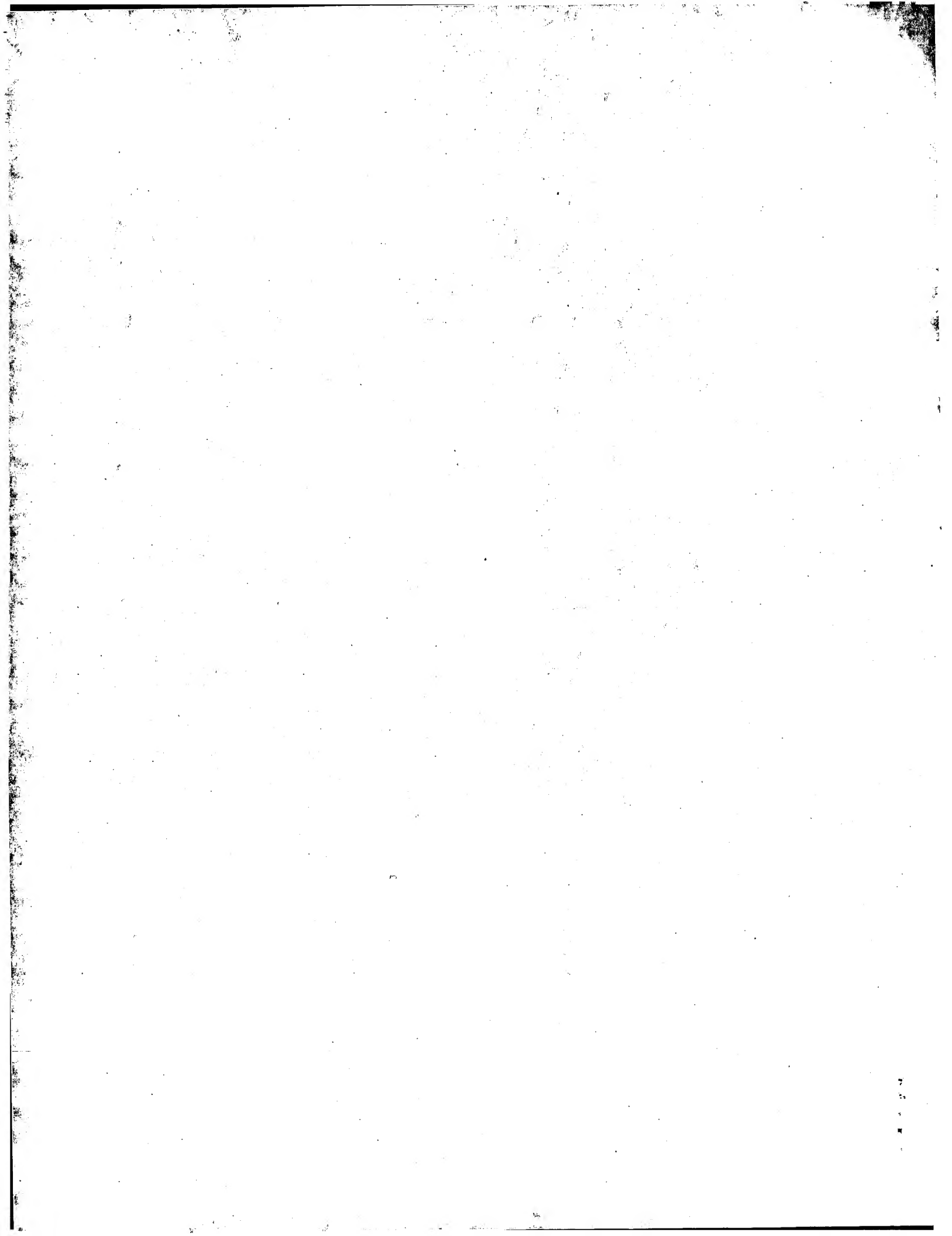
ACCESSION BQ619318
VERSION BQ619318.1 GI:21621312
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Wang, H. and Bohnert, H.J.
AUTHORS Genomics of plant stress tolerance
TITLE Unpublished (2002)

FEATURES Email: bohnertlab@life.uiuc.edu.
source Location/Qualifiers
1. .648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6E12_SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619337 (1-648)
QY 1 MetAlaTyrTyrGlnGluValAlaAspTyrCysSerGluValArgSerValAlaProAla 20
Db 82 ATGGCTTACTACCGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 142 GGCTTCGGCGCCACGAGCGGCGTCCAGCAGCACCGTCTGCAAGGAGAAAGTTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisHisGly 59
Db 202 GTCGACACGGTCTCACGCGCGCGGCCAACCCACCACCACCACCATGTCACCGCGC 261
QY 60 GlyHisGlyPheValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCGTGGTGGCGGAGACCAGGGTCCGAGGAGGACATCACACCTGCACCGGC 321
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGCAGGAGAGAGCTTCCTCGCCAGGCGCTAAC 363

Search completed: June 16, 2003, 12:16:33
Job time : 1303.5 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 04:03:20 ; Search time 1684.8 Seconds
(without alignments)
9915.101 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgcacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```
GenEmbl:*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.mu.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em_sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	44	7.7	173699	8	AP003416
C 2	44	7.7	174218	2	AC091071
C 3	44	7.7	185095	8	AP003436
C 4	44	7.7	185481	2	AC091088
C 5	31	5.4	185133	2	AC119716
C 6	29	5.1	569	6	AX400832
C 7	29	5.1	1138	9	BC032756
C 8	29	5.1	1275	8	AY102542
C 9	29	5.1	1450	5	AF295407
C 10	29	5.1	2879	10	BC012959
C 11	29	5.1	3194	9	AK025519
C 12	29	5.1	30523	2	AC115592
C 13	29	5.1	112226	9	AC099486
C 14	29	5.1	172823	9	AC114323
C 15	28	4.9	410	8	AF467161
C 16	28	4.9	411	8	AF467129
C 17	28	4.9	411	8	AF467148
C 18	28	4.9	411	8	AF467156
C 19	28	4.9	881	3	AY044834
C 20	28	4.9	1436	9	BC015065
C 21	28	4.9	1635	9	BC010519
C 22	28	4.9	1849	9	BC008675
C 23	28	4.9	1853	3	AB062683
C 24	28	4.9	1865	9	BC012562
C 25	28	4.9	2115	10	BC007136
C 26	28	4.9	5928	6	AX346988
C 27	28	4.9	6593	6	AX345380
C 28	28	4.9	13919	6	AX356480
C 29	28	4.9	42499	8	AC004625
C 30	28	4.9	77722	9	AC093762
C 31	28	4.9	100652	2	AC116098
C 32	28	4.9	130010	2	AC025292
C 33	28	4.9	132000	2	AC116976
C 34	28	4.9	133888	2	AC010769
C 35	28	4.9	152472	2	AP000486
C 36	28	4.9	155354	2	AC107591
C 37	28	4.9	159658	9	AC104062
C 38	28	4.9	168145	9	AL159171
C 39	28	4.9	170022	2	AC021559
C 40	28	4.9	176827	9	AC022306
C 41	28	4.9	179191	2	AL808114
C 42	28	4.9	209764	10	AL645468
C 43	28	4.9	216250	9	AC099731
C 44	27	4.7	330	11	G41782
C 45	27	4.7	369	3	AF400199

ALIGNMENTS

RESULT 1
AP003416/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0456E05.
ACCESSION
AP003416
VERSION
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,
clone: P0456E05.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1

CDS

complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/gene="P0456E05.7")
/note="contains ESTs AU101665(R2414),D24709(R2594)
similar to Arabidopsis thaliana chromosome 5, At5g14420
unknown protein"
/codon_start=1
/protein_id="BAB92608.1"
/db_xref="GI:20804929"
/translation="MGQGSRSRDRGSGHRYGSHPSFQQQWGGGGGGGGGGGYPY
GDADHGQYVQAPQGGYAAAPYQPPMPSPAATOPPARAGGAKPRLDRRYS
RIADYHSYQVTDALAQAGLESSNLIVGIDTKSNEWLGKSEHGRSLHHISNAPN
YQAHSIIIGOTLSKEDENLIPCFGSDATHDQVFCYFDLRPCNGCSALSYRE
LVPRLRAGTSFAPLIEMANTIVEQSGQVHLIIADQGVTSVDTASQLSQEQ
KVDIAVRASELPISLVIVGVGDPWDMKEFDNIPRAFDFQVNFSEIMKNMP
QSRKKAASLMEIPQYKATVELGILGRSFKSPERVPLPPGGSHDAYSYGSKS
FKPSYFQSSSTSSPHYETAQSSPAVPSSTVDNQVCPICLVNPKDMAFGGCHQC
NPKAFERLADKPKVYWCIDKDRKASNPQLQDCGAGDVSSIETH"
51367..52887
/gene="P0456E05.8"
51367..52887
/gene="P0456E05.8"
/note="hypothetical protein
similar to salt-inducible protein"
/codon_start=1
/protein_id="BAB92609.1"
/db_xref="GI:20804930"
/translation="MAASRLRLKPLSLISKQRLSPIDVEQAESPASSSIPID
PSLPIPLAVSHLSPSPALPSPVHASTPAALLRIILRRHHPRLAALDLHLLAA
SDSAFRPDRLTSLASRAESRLPLSLRLLEVLFRPCPTDSDIFACPELLTF
RKAILAASGDIPAAADALASRAADSPAPAEFYNIILHALARLHEDTVFYGE
MTSTRHVPDYTTNIIILNSCRAGVDAAWRFOEMRWSCTSPGVSNFLMRGFR
BGRCKEGLKVAEMQLQGLSVASMEIINGLRCGCEPKAAEYFEFLDGVVPEG
FDCLDLVESLCVRNVEKAVVVELILERNWVSLGVPAQVTECLMKEGKLDKACQ
MGMWRAAEIVDPISNYIFALCEAGRTVDANRLQAKRKGFOADGFTYSIVQG
FGROIRKEGAEVLDEMDAGIVPNIAITNRLDLHMRGSMQLOKQCSRHNAAN"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9")

gene

Query Match 7.7%; Score 44; DB 8; Length 173699;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS

QY 103 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 146
Db 157564 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 157521

Query Match

Best Local Similarity 7.7%; Score 44; DB 8; Length 173699;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS

AC091071 174218 bp DNA linear HTG 27-MAR-2001
DEFINITION Oriza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

ACCESSION

AC091071
VERSION HTG_13450004
KEYWORDS HTG: HTGS_PHASE2.

SOURCE

ORGANISM Oriza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartodeae; Oryzaceae; Oryza.
1 (bases 1 to 174218)
Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
and Eun, M.Y.

REFERENCE

Unpublished
TITLE Oriza sativa PAC P0456E05 genomic sequence
JOURNAL
REFERENCE 2 (bases 1 to 174218)
Hahn, J.-H., Eun, M.Y. and Kim, H.-I.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Rice Genome Sequencing Project, National

COMMENT

Institute of Agricultural Science and Technology (NIAT), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail:jhahn@rda.go.kr,
Tel:82-31-290-0309, Fax:82-31-290-0308)
The PAC clone was provided by Japanese Rice Genome Research Program
of NIAR.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 53113: contig of 53113 bp in length
* 53114 53213: gap of unknown length
* 53214 172404: contig of 119191 bp in length
* 172405 172504: gap of unknown length
* 172505 174218: contig of 1714 bp in length.
FEATURES
Location/Qualifiers
1..174218
/organism="Oriza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0456E05"
BASE COUNT 49037 a 38377 c 37185 g 49417 t 202 others
ORIGIN
Query Match 7.7%; Score 44; DB 2; Length 174218;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 146
Db 16068 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 16111
RESULT 3
AP003436/c 185095 bp DNA linear PLN 17-APR-2002
LOCUS Oriza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION PAC clone:P0470A12.
ACCESSION AP003436
VERSION AP003436.2 GI:16197551
KEYWORDS Oriza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0470A12.
ORGANISM Oriza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartodeae; Oryzaceae; Oryza.
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oriza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12
Published Only in Database (2001)
2 (bases 1 to 185095)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (21-MAR-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Oct 17, 2001 this sequence version replaced gi:13430001.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGF. Protein homologues of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent


```

AARISGVKSVDSYDDVATRLKAVIKDNTLSLELYSVGGLLSIKRQHNVLPLD
ADNTFHAIKALSODGRWYDNTSAESSTFAAGIALEALSASVDSGSMIAV
KNDIVLFDYIKSYDDGTFFDEKHVDAAYKPIITTSASVVRGVTSPRAVAGKINI
PGKILGLAKFFLIGLPGSKDCFNQIEVFLVLSLPSKVFSTSKDLKVEVTTV
FGSAAPPLRVNLVQVLGSDSKVITTEKLOPDLNNVHYLDIAPLKIDVGYLSVFE
ISLQEQEHETIYATGTNTAEAFVETGLIKVDAEIGISDAGTVESVOKIDLQKDTL
VSLSANHLKRLSFLSPLGKTFKPFPHOVFLKLKHDESKVEHLFVPGSARQFQIVL
DFGLVKEVYVLSGRYDLDELAVDAAMENSFLRALGHIELDLPEAPKAPKPPAQAVD
PESKFPKKEISHIFSPKRPKLSFAFTGLTLLPIVGLFGLMRGLVKNLKNFPLS
PAPAFASFLHAGIGAVLLYVLFWIKLDTFTLKYLSPLGLVFLVFGHRLSYLSST
SAKQKTA"
complement(Join(47978)..48061,48179..48239,48314..48378,
48478..48759,48837..48911,49020..49094,49242..49298,
49552..49625,49763..49964,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52039,52148..52561))
/clone="P0470A12.8"
complement(Join(47978)..48061,48179..48239,48314..48378,
48478..48759,48837..48911,49020..49094,49242..49298,
49552..49625,49763..49964,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52039,52148..52561))

Query Match
Best Local Similarity 7.7%; Score 44; DB 8; Length 185095;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACGAGGAGTGGACTACTCTCGGAGGAGTGGAGTGGTG 146
|||||
DB 27507 ACTACGAGGAGTGGACTACTCTCGGAGGAGTGGAGTGGTG 27464
|||||

AC091088 185481 bp DNA linear HTG 29-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0470A12,
*** SEQUENCING IN PROGRESS ***; 4 ordered pieces.
AC091088
AC091088.1 GI:13487944
HTG: HTGS_PHASE2.
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 185481)
Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
and Eun, M.-Y.
Oryza sativa PAC P0470A12 genomic sequence
2 (bases 1 to 185481)
Unpublished
Hahn, J.-H., Eun, M.-Y. and Kim, H.-I.
Direct Submission
Submitted (29-MAR-2001) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology (NIAT), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail:jhahn@rda.go.kr,
Tel:82-31-290-0309, Fax:82-31-290-0308)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given. However, the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11215: contig of 11215 bp in length
* 11316 11315: gap of unknown length
* 93352 93352: contig of 82037 bp in length
* 93353 93452: gap of unknown length
* 93453 138783: contig of 45331 bp in length
* 138784 138883: gap of unknown length
* 138884 185481: contig of 46598 bp in length.
Location/Qualifiers
FEATURES

```

source

```

1..185481
/organism="Oryza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0470A12"
BASE COUNT 51439 a 41522 c 41493 g 50653 t 374 others
ORIGIN

```

Query Match

```

Best Local Similarity 7.7%; Score 44; DB 2; Length 185481;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 103 ACTACGAGGAGTGGACTACTCTCGGAGGAGTGGAGTGGTG 146
|||||
DB 156728 ACTACGAGGAGTGGACTACTCTCGGAGGAGTGGAGTGGTG 156771
|||||

```

RESULT 5

```

AC119716
LOCUS

```

```

DEFINITION

```

```

AC119716

```

```

ACCESSION

```

```

VERSION

```

```

KEYWORDS

```

```

SOURCE

```

```

ORGANISM

```

REFERENCE

```

AUTHORS

```

```

1 (bases 1 to 185133)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,
Alsbrooks, S.L., Anaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Delaney, K.R., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Dearth, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Eathart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flaggs, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, J., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H.,
Lopez, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Tellford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission

```

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 185133)
Worley K.C.
Direct Submission
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 185133)
Worley K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 18, 2002 this sequence version replaced gi:20467917.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWLC
Center clone name: CH230-274116
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 120616 bases at least Q40
Consensus quality: 127908 bases at least Q30
Consensus quality: 132518 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 77 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1001: contig of 1001 bp in length
1002
1101: gap of unknown length
1102: contig of 1037 bp in length
2139
2238: gap of unknown length
2239
3632: contig of 1394 bp in length
3633
3732: gap of unknown length
3733
4771: contig of 1039 bp in length
4772
4871: gap of unknown length
6244: contig of 1373 bp in length
6245
6344: gap of unknown length
6345
7418: contig of 1074 bp in length
7419
7518: gap of unknown length
8565: contig of 1047 bp in length
8566
8665: gap of unknown length
8666
9554: contig of 1189 bp in length
9555
11359: contig of 1405 bp in length
11360
12534: contig of 1075 bp in length
12535
12634: gap of unknown length
12635
13718: contig of 1084 bp in length
13719
13818: gap of unknown length
13819
14940: contig of 1122 bp in length
14941
15040: gap of unknown length
15041
16428: contig of 1388 bp in length
16429
16528: gap of unknown length
16529
17844: contig of 1316 bp in length
17845
19405: gap of unknown length
19406
19205: gap of unknown length
20676: contig of 1471 bp in length
20677
20776: gap of unknown length
22169: contig of 1393 bp in length
22777

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

22170
22270
22270
23379
23379
24551: gap of unknown length
24552
24552
25824: contig of 1173 bp in length
25825
25924: gap of unknown length
25925
27234: contig of 1310 bp in length
27235
27334: gap of unknown length
27335
27335
28742: contig of 1408 bp in length
28743
28842: gap of unknown length
30052: contig of 1210 bp in length
30053
30152: gap of unknown length
31262: contig of 1110 bp in length
31263
31362: gap of unknown length
31363
32586: contig of 1224 bp in length
32587
32686: gap of unknown length
34161: contig of 1475 bp in length
34162
34261: gap of unknown length
34262
36688: contig of 2427 bp in length
36689
36788: gap of unknown length
38026: contig of 1238 bp in length
38027
38126: gap of unknown length
38127
39339: gap of unknown length
39340
41317: contig of 1378 bp in length
41318
41417: gap of unknown length
43191: contig of 1774 bp in length
43192
43291: gap of unknown length
43292
44826: contig of 1535 bp in length
44827
44926: gap of unknown length
46264: contig of 1338 bp in length
46265
46364: gap of unknown length
48154: contig of 1790 bp in length
48155
48254: gap of unknown length
48255
50357: contig of 2103 bp in length
50358
50457: gap of unknown length
50458
51775: contig of 1318 bp in length
51776
51875: gap of unknown length
51876
53265: contig of 1390 bp in length
53266
53365: gap of unknown length
53366
55987: contig of 2622 bp in length
55988
56087: gap of unknown length
58175: contig of 2088 bp in length
58176
58275: gap of unknown length
58276
60285: contig of 2010 bp in length
60286
60385: gap of unknown length
60386
62842: contig of 2257 bp in length
62843
62742: gap of unknown length
62743
64127
64226: gap of unknown length
64227
65801: contig of 1575 bp in length
65802
65901: gap of unknown length
65902
68814: contig of 2913 bp in length
68815
68914: gap of unknown length
70953: contig of 2039 bp in length
70954
71053: gap of unknown length
73492: contig of 2439 bp in length
73493
73592: gap of unknown length
73593
74956: contig of 1364 bp in length
74957
75056: gap of unknown length
75057
76926: contig of 1870 bp in length
76927
77026: gap of unknown length
77027
79405: contig of 2379 bp in length
79406
79505: gap of unknown length
80966: contig of 1461 bp in length
80967
81066: gap of unknown length
81067
83641: contig of 2575 bp in length
83642
83741: gap of unknown length
83742
87177: contig of 3436 bp in length
87178
87277: gap of unknown length
87278
89743: contig of 2466 bp in length
89744
89843: gap of unknown length

```

*      89844      91514: contig of 1671 bp in length

Query Match
Best Local Similarity 5.4%; Score 31; DB 2; Length 185133;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGCACAGC 31
|||||
Db 51879 ACCACGCGTCCGCCACGCGTCCGCACAGC 51909
|||||

RESULT 6
AX400832/c
LOCUS AX400832 569 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 508 from Patent WO0210453.
ACCESSION AX400832
VERSION AX400832.1 GI:21337012
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Eliashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 508 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
Source Location/Qualifiers
1..569
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="PMBL/GenBank Accession No. AI007824"
BASE COUNT 132 a 98 c 117 g 222 t
ORIGIN
Query Match
Best Local Similarity 5.1%; Score 29; DB 6; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAATACTAAAAAATAAAAAAAAAAAAAA 574
|||||
Db 73 GTAATACTAAAAAATAAAAAAAAAAAAAA 45
|||||

RESULT 7
BC032756
LOCUS BC032756 1138 bp mRNA linear PRI 26-JUN-2002
DEFINITION Homo sapiens, H2A histone family, member C, clone MGC:44857
IMAGE:5581631, mRNA, complete cds.
ACCESSION BC032756
VERSION BC032756.1 GI:21595675
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NTH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL),
National Institutes of Health Intramural
Sequencing Center (NISC),

Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breese, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 69 Row: p Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15718718.

FEATURES
Source Location/Qualifiers
1..1138
/organism="Homo sapiens"
/db_xref="LocusID:8329"
/db_xref="taxon:9606"
/clone="MGC:44857 IMAGE:5581631"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_67"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
52..444
/codon_start=1
/product="H2A histone family, member C"
/protein_id="AAH32756.1"
/db_xref="GI:21595676"
/translation="MSGKQGGKGGKARAKTSSRAGLQFPVGRVHLLRGNVAERV
GAGPVLAAYLVLTAEITLAEAGNADNKKTRIPRLQLAIRNDEELNKLGLKVT
IAQGVLPNIQAVLLPKKTESHHKANG"
BASE COUNT 312 a 315 c 303 g 208 t
ORIGIN
Query Match
Best Local Similarity 5.1%; Score 29; DB 9; Length 1138;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCACGCGTCCGCCACGCGTCCGCACAG 30
|||||
Db 12 CCCACGCGTCCGCCACGCGTCCGCACAG 40
|||||

RESULT 8
AY102542
LOCUS AY102542 1275 bp mRNA linear PLN 13-AUG-2002
DEFINITION Arabidopsis thaliana hypothetical protein (At2g23370/F2686.2) mRNA,
complete cds.
ACCESSION AY102542
VERSION AY102542.1 GI:21805664
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1275)
Xiao, Y., Ishmael, N., Kumar, N., Redman, J., Riedmuller, S.,
Utterback, T., Whitelaw, C.A., Fraser, C.M. and Town, C.D.
TITLE Cloning and sequencing of full-length cDNAs for hypothetical genes
from chromosome 2 of Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 1275)
Xiao, Y., Ishmael, N., Kumar, N., Redman, J., Riedmuller, S.,
Utterback, T., Whitelaw, C.A., Fraser, C.M. and Town, C.D.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-2002) Plant Genomics, The Institute for Genomic
Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers

```

```
1. .1275
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="2"
/notes="ecotype: Columbia"
1. .1275
/gene="At2g23370/F26B6.2"
29. .1051
/gene="At2g23370/F26B6.2"
/codon_start=1
/evidence="not_experimental"
/product="hypothetical protein"
/protein_id="AA076747.1"
/db_xref="GI:21805665"
/translation="MHRGAAGVGRKRRLLHGRNFAKSCSSSLSEHCEFKINDIA
KAQPKHMLNLIITARGKSTVSPCAKQGLLPIPLVKMSPGSSIALLRWPTAPSS
MEMPVVEVQKHGWFLANNVDQFIHRIIVEDVSKPECSQEIFNAAGEAGKKLYSKG
DFASSRLMDLALYLLKRVGLFSDSLERKVIHRIENGDRHVSALVATEFTKRGNPFQFA
RPFANAKVLLKLRNLEAKDAARGALKSSWTLCGRVEETIAQIAEWGEEOIAQYKER
VTGEGQRDIDRGPMAQASLDEAAFLNLASLEGTWDESLERVAQCYKEAGLNDIAK
FVLYRD"
1183
/gene="At2g23370/F26B6.2"
/notes="alternative"
1240
/gene="At2g23370/F26B6.2"
/notes="alternative"
1251
/gene="At2g23370/F26B6.2"
386 a 232 c 308 g 348 t 1 others
BASE COUNT
ORIGIN
Query Match 5.1%; Score 29; DB 8; Length 1275;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 546 GTAATACTAAAAA...AAAAA 574
|||||
DB 1242 GTAATACTAAAAA...AAAAA 1270
|||||
RESULT 9
AF295407 1450 bp mRNA linear VRT 01-SEP-2001
LOCUS
DEFINITION
Danio rerio alcohol dehydrogenase (ADH) mRNA, complete cds.
ACCESSION
AF295407
VERSION
AF295407.1 GI:15428577
KEYWORDS
Danio rerio.
SOURCE
Danio rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
Tanguay,R.L. and Vasillou,V.
1 (bases 1 to 1450)
Identification and characterization of the zebrafish alcohol
dehydrogenase
unpublished
JOURNAL
REFERENCE
2 (bases 1 to 1450)
Tanguay,R.L. and Vasillou,V.
Direct Submission
TITLE
Submitted (15-AUG-2000) School of Pharmacy, University of Colorado
Health Sciences Center, 4200 East Ninth Ave, Box C238, Denver, CO
80262, USA
LOCATION/Qualifiers
FEATURES
source
1. .1450
/organism="Danio rerio"
/db_xref="taxon:7955"
1. .1450
/gene="ADH"
49. .1182
/gene="ADH"
/codon_start=1
gene
CDS
1. .1275
/product="alcohol dehydrogenase"
/protein_id="AAK97853.1"
/db_xref="GI:15428578"
/translation="MAFACKVIKRAAVAKPAKLMMEIEIVAPQEGEIRKVIAT
GLCHTDLYHLVDGDKRGKFPVILGHESAGVVGCVTDYKPGDKVPLPLSQCGKCK
FCCKPTNLCESSWATKYHDIMAEPTSRCTRGQTILOPMGTSTSEYTVINONAVAK
IDENAPLDRLVFLGCGITTYGAAVNTAGTIPGSCAVFGIAGVGLAAVYMGCKNAGAS
RIFADVINEKKFKAKVFGVDFLNPFAFNKPISEVLIEMTGGVGVDFSECTGNEIV
MRSALESCAKMGVSVVVGWTVNDFSAKPIQLIYGTWKGLSFGGFKCKDSVPKLVLR
DYMGGKIMLDEFITKMLNLEQVNDALNLMKTGOGIRTINTVSK"
419 a 316 c 356 g 359 t
BASE COUNT
ORIGIN
Query Match 5.1%; Score 29; DB 5; Length 1450;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CCACGGTCGCCGCCACGCGTCCGCACAGC 31
|||||
DB 1 CCACGGTCGCCGCCACGCGTCCGCACAGC 29
|||||
RESULT 10
BC012959 2879 bp mRNA linear ROD 07-AUG-2002
LOCUS
DEFINITION
Mus musculus, staufer (RNA-binding protein) homolog 1 (Drosophila),
clone MGC:13708 IMAGE:4188394, mRNA, complete cds.
ACCESSION
BC012959
VERSION
BC012959.1 GI:15277959
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2879)
Strausberg,R.
Direct Submission
Submitted (20-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 18 Row: j Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755673.
LOCATION/Qualifiers
FEATURES
source
1. .2879
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:13708 IMAGE:4188394"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
290. .1759
/codon_start=1
/product="staufer (RNA-binding protein) homolog 1
(Drosophila)"
/protein_id="AAH12959.1"
CDS
```

```

/db_xref="GI:15277960"
/db_xref="LocusID:20853"
/translation="MYKVPDHSRQSTYSYGMRGAGVPPRYFFPFPVPLLYQVELS
VGGQNPNGKGRPPVKHDAPARALTQSEPLPERLEVNGREAEENLNKSEISQVF
ETALKNLVNEFEARESGPPHMKNEVTVSYGFEVGEKSKISKNNAARVLEQ
LRRLPAPARFVKPRIRKKSOPTCKLATPDYGGMNPISRLAQIOAKKEPEYM
LITERGLPRRFRVQVGVGHHTAGVGTNKVKARNAENMLETLGPKVQAOQAPK
ALKSEKTPVKPGDKQKVTFPEPSGDCGTSNKDEEFMPYLSHQQLPAGILPWVP
EVAQAVGVSGHHTKDFPAPNPAPAKATVTAMIARELLYGTSTAEITLILNSISGH
VPHGPRTRSEGLYLSRAQGVQVEIKDFPPKNNKCVSLNCSQPPFLVSHGIGKDV
ESCHDMAALNLIKLSLSELDQOSTEMPRGTGPGVPSACGR"
BASE COUNT      798 a 735 c 699 g 647 t
ORIGIN

Query Match      5.1%; Score 29; DB 10; Length 2879;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAACTACTAAAAA 574
|||||
Db 2821 GTAATACTAAAAA 2849
|||||

RESULT 11
AK025519 3194 bp mRNA linear PRI 29-SEP-2000
LOCUS
DEFINITION
Homo sapiens cDNA: FLJ21866 fis, clone HEP02379, highly similar to
AF061938 Homo sapiens staufen protein (STAU) mRNA.
ACCESSION
AK025519
VERSION
AK025519.1 GI:10438059
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens Hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
clone:HEP02379.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 3194)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEBO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction; 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1. .3194
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP02379"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"
misc_feature
1. .3194
/notes="highly similar to AF061938 Homo sapiens staufen
protein (STAU) mRNA"
BASE COUNT      894 a 787 c 732 g 781 t
ORIGIN

Query Match      5.1%; Score 29; DB 9; Length 3194;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAACTACTAAAAA 574
|||||
Db 3165 GTAATACTAAAAA 3193
|||||

RESULT 12
AC115592 30523 bp DNA linear HTG 21-MAR-2002
LOCUS
DEFINITION
Dictyostelium discoideum chromosome 2 map 109820-140341 strain AX4,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC115592
VERSION
AC115592.1 GI:19569975
KEYWORDS
HTG; HTGS PHASE2.
SOURCE
Dictyostelium discoideum.
ORGANISM
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
REFERENCE
1 (bases 1 to 30523)
Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
Noegel,A.A.
Sequence and Analysis of Chromosome 2 of Dictyostelium
Unpublished
The Dictyostelium Genome Sequencing Consortium
2 (bases 1 to 30523)
Baumgart,C.
Direct Submission
Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
CDS predictions from GeneID may contain errors. Further Information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.unl-koein.de/dictyostelium/project.shtml
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1. .30523
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/chromosomes="2"
/map="109820-140341"
697. .971
/notes="ORF_ID:dd_01553"
/pseudo
/codon_start=1
1886. .2778
/notes="ORF_ID:dd_01554"
/codon_start=1
/product="40S ribosomal protein S15A"
/protein_id="AAL92278.1"
/db_xref="GI:19569976"
/translation="MYRISLVNCLNLTVAEROCKROVLVPPSSKVIKLEVMKK
RYIGFEIVDDHRSKIVIDLGRINKGVISPRDVLDEIEKWASLLPSRQFGHI
VLTSIGIMDHNEAKTRHTGKLLGFFY
complement(join(2304. .2668,2759. .3450,3569. .3911))
/notes="ORF_ID:dd_01555"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAL92279.1"
/db_xref="GI:19569977"
/translation="MINELISKFKKVIYGFNCLVGEIFKSNKHLIGTVGDKII
VYLNDDEIEVLETKGSSIQLELLDTTKFGSDVNSGNLVFSGHQLYRDTLN
GSITSITRLSIGSLGLVTSIKPHQNAIWRKIPSVNQTLSLIDSLIDAINPNIK

```



```

RESULT 13
AC099486
LOCUS 112226 bp DNA linear PRI 01-JAN-2002
DEFINITION Homo sapiens chromosome 5 clone CTD-2011G10, complete sequence.
ACCESSION AC099486
VERSION AC099486.2 GI:18030138
SOURCE HTG.
ORGANISM Homo sapiens.
            Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 112226)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
AUTHORS 2 (bases 1 to 112226)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL
AUTHORS 2 (bases 1 to 112226)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL
COMMENT Submitted (01-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 1, 2002 this sequence version replaced gi:16930902.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
FEATURES
source
1. 112226
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2011G10"
BASE COUNT 36508 a 19376 c 19390 g 36952 t
ORIGIN
Query Match 5.1%; Score 29; DB 9; Length 112226;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 574
|||||
Db 90165 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 90193
|||||

RESULT 14
AC114323/c
LOCUS 172823 bp DNA linear PRI 23-APR-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-93N19, complete sequence.
ACCESSION AC114323
VERSION AC114323.2 GI:20270116
SOURCE HTG.
ORGANISM Homo sapiens.
            Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 172823)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
AUTHORS 2 (bases 1 to 172823)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL
COMMENT Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 172823)

```

```

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 23, 2002 this sequence version replaced gi:19224972.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.3.
FEATURES
source
1. 172823
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-93N19"
BASE COUNT 56281 a 29916 c 30420 g 56206 t
ORIGIN
Query Match 5.1%; Score 29; DB 9; Length 172823;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 574
|||||
Db 166787 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 166759
|||||

RESULT 15
AF467161
LOCUS 410 bp DNA linear PLN 15-JUL-2002
DEFINITION Lasthenia gracilis isolate calli153 trna-Lys (trnK) gene, intron;
and maturase K (matK) gene, partial cds; Chloroplast genes for
chloroplast products.
ACCESSION AF467161
VERSION AF467161.1 GI:21759599
KEYWORDS
SOURCE
ORGANISM Lasthenia gracilis.
            Chloroplast Lasthenia gracilis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
            Helenieae; Lasthenia.
            1 (bases 1 to 410)
            Chan,R., Baldwin,B.G. and Ornduff,R.
            Cryptic goldfields: A molecular phylogenetic re-investigation of
            Lasthenia californica sensu lato and close relatives (Compositae:
            Heliantheae sensu lato)
            Am. J. Bot. 89 (7), 1103-1112 (2002)
            2 (bases 1 to 410)
            Chan,R., Baldwin,B.G. and Ornduff,R.
            Direct Submission
            Submitted (10-JAN-2002) Jepson Herbarium and Department of
            Integrative Biology, University of California, 1001 Valley Life
            Sciences Building # 2465, Berkeley, CA 94720-2465, USA
FEATURES
source
1. 410
/organism="Lasthenia gracilis"
/organelle="plastid:chloroplast"
/isolate="calli153"
/specimen_voucher="Neese and Painter 2552 (JEPS)"
/db_xref="taxon:182806"
<1..>410
/gene="trnK"
/note="trna-Lys"
<1..>410
/gene="trnK"
<1..175
/gene="matK"
<1..175
/gene="matK"
/codon_start=2

```

```
/transl_table=11
/product="maturase K"
/protein_id="AA077460.1"
/db_xref="GI:21759600"
/translation="RAFLKRFQSLLLEFFTEEEQVFSLTFFRVSSISRRLSRRIWY
LDIVCINDLANVE"
BASE COUNT      139 a      48 c      81 g      142 t
ORIGIN
Query Match      4.9%; Score 28; DB 8; Length 410;
Best Local Similarity 100.0%; Pred. NO. 0.00075;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      547 TAATACTAAAAA
Db      221 TAATACTAAAAA
```

Search completed: June 16, 2003, 07:56:05
Job time : 1698.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 06:35:31 ; Search time 39.9504 Seconds
(without alignments)
4406.275 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 acccagcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/lna/backfilesi.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	4.5	212	4	US-09-280-116-50
2	26	4.5	638	4	US-08-998-416-981
3	26	4.5	717	4	US-08-998-416-981
4	26	4.5	792	3	US-08-950-720A-1
5	26	4.5	1021	4	US-03-280-116-85
6	26	4.5	1210	4	US-03-560-639-12
7	26	4.5	1384	4	US-09-372-422A-17
8	26	4.5	1448	2	US-08-942-218A-1
9	26	4.5	1504	4	US-09-280-116-1
10	26	4.5	2205	3	US-08-888-077A-41
11	26	4.5	2864	4	US-09-409-180A-2
12	26	4.5	3416	2	US-08-724-394A-15
13	26	4.5	3747	2	US-08-044-618-5
14	26	4.5	4880	3	US-09-031-563-1
15	26	4.5	4880	4	US-09-392-277-1
16	26	4.5	8519	4	US-09-261-907-1
17	26	4.5	162450	4	US-09-345-882-1
18	25	4.4	140	1	US-08-628-417-5
19	25	4.4	216	1	US-08-686-878A-34
20	25	4.4	216	4	US-09-175-928-34
21	25	4.4	240	1	US-08-628-417-6
22	25	4.4	330	4	US-09-078-294-24
23	25	4.4	493	4	US-09-280-116-166
24	25	4.4	493	4	US-09-392-184-20
25	25	4.4	508	4	US-09-280-116-47
26	25	4.4	601	4	US-09-336-536-74
27	25	4.4	728	4	US-09-091-097-5

```

28 25 4.4 749 4 US-09-257-583-12 Sequence 12, Appl
29 25 4.4 810 4 US-09-072-993C-5 Sequence 5, Appl
30 25 4.4 944 4 US-09-227-357-122 Sequence 122, App
31 25 4.4 1132 3 US-08-894-731-3 Sequence 3, Appl
32 25 4.4 1148 4 US-09-395-674B-7 Sequence 7, Appl
33 25 4.4 1164 2 US-08-794-796-1 Sequence 1, Appl
34 25 4.4 1284 4 US-09-286-529-6 Sequence 6, Appl
35 25 4.4 1353 4 US-08-913-014A-6 Sequence 44, Appl
36 25 4.4 1376 4 US-09-443-184-44 Sequence 4, Appl
37 25 4.4 1454 4 US-09-372-422A-19 Sequence 19, Appl
38 25 4.4 1619 2 US-08-991-946A-2 Sequence 2, Appl
39 25 4.4 1930 4 US-08-987-367-1 Sequence 7, Appl
40 25 4.4 2181 4 US-09-254-465A-7 Sequence 11, Appl
41 25 4.4 2181 4 US-09-254-465A-11 Sequence 23, Appl
42 25 4.4 2191 4 US-09-399-913-25 Sequence 25, Appl
43 25 4.4 2191 4 US-09-298-731-25 Sequence 3, Appl
44 25 4.4 2197 1 US-08-233-005-3 Sequence 3, Appl
45 25 4.4 2197 1 US-08-428-943-3

```

ALIGNMENTS

```

RESULT 1
US-09-280-116-50
; Sequence 50, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: zinc carboxypeptidases
US-09-280-116-50

```

```

Query Match 4.5%; Score 26; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ACCCAGCGTCGCCGCCACGCGTCCGC 26
Db 19 ACCCAGCGTCGCCGCCACGCGTCCGC 44

```

```

RESULT 2
US-08-998-416-981/c
; Sequence 981, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6239264artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA

```


;
; NAME/KEY: Coding Sequence
; LOCATION: 77...481
; OTHER INFORMATION:
US-08-950-720A-1

Query Match 4.5%; Score 26; DB 3; Length 792;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGGCTCCGC 26
|||||
DB 42 ACCACGGCTCCGCCACGGCTCCGC 17

RESULT 5

US-09-280-116-85/c
; Sequence 85, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: aspartyl proteases
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1021)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-85

Query Match 4.5%; Score 26; DB 4; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGGCTCCGC 26
|||||
DB 38 ACCACGGCTCCGCCACGGCTCCGC 13

RESULT 6

US-09-560-639-12
; Sequence 12, Application US/09560639
; Patent No. 6323334
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 7853-158
; CURRENT APPLICATION NUMBER: US/09/560,639
; CURRENT FILING DATE: 2000-04-28
; EARLIER APPLICATION NUMBER: 60/155,862
; EARLIER FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(557)
US-09-560-639-12

Query Match 4.5%; Score 26; DB 4; Length 1210;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGGCTCCGC 26
|||||
DB 5 ACCACGGCTCCGCCACGGCTCCGC 30

RESULT 7

US-09-372-422A-17
; Sequence 17, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)...(959)
US-09-372-422A-17

Query Match 4.5%; Score 26; DB 4; Length 1384;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGGCTCCGC 26
|||||
DB 10 ACCACGGCTCCGCCACGGCTCCGC 35

RESULT 8

US-08-942-218A-1/c
; Sequence 1, Application US/08942218A
; Patent No. 5935835
; GENERAL INFORMATION:
; APPLICANT: McCard-Roshak, Amy
; APPLICANT: Marshall, Lisa
; TITLE OF INVENTION: Human Myt-1 Kinase Clone
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/942,218A
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,389
; FILING DATE: 11-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031

```

; REFERENCE/DOCKET NUMBER: ATG-50027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-942-218A-1

Query Match
Best Local Similarity 4.5%; Score 26; DB 2; Length 1448;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
Db 46 ACCACGCGTCCGCCACGCGTCCGC 21

RESULT 9
US-09-280-116-1/C
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
; US-09-280-116-1

Query Match
Best Local Similarity 4.5%; Score 26; DB 4; Length 1504;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
Db 64 ACCACGCGTCCGCCACGCGTCCGC 39

RESULT 10
US-08-888-077A-41
; Sequence 41, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888.077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2205
; OTHER INFORMATION: /note= "mutTM1-TM2"
; US-08-888-077A-41

Query Match
Best Local Similarity 4.5%; Score 26; DB 3; Length 2205;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
Db 15 ACCACGCGTCCGCCACGCGTCCGC 40

RESULT 11
US-09-409-180A-2
; Sequence 2, Application US/09409180A
; Patent No. 6444802
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: White, David
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 22196, A No. 644802el Human Aminopeptidase
; FILE REFERENCE: 5800-59
; CURRENT APPLICATION NUMBER: US/09/409,180A
; CURRENT FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (300)...(2414)
; US-09-409-180A-2

Query Match
Best Local Similarity 4.5%; Score 26; DB 4; Length 2864;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
Db 250 ACCACGCGTCCGCCACGCGTCCGC 275

RESULT 12
US-08-724-394A-15/C
; Sequence 15, Application US/08724394A
; Patent No. 5872237
```

```

; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..3416
; OTHER INFORMATION: /note= "cDNA 44"
; US-08-724-394A-15

```

```

Query Match 4.5%; Score 26; DB 2; Length 3416;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAGCGTCCGCCACGCGTCCGC 26
Db 63 ACCCAGCGTCCGCCACGCGTCCGC 38

RESULT 13
US-08-044-618-5
; Sequence 5, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654, 0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3747 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-044-618-5

Query Match 4.5%; Score 26; DB 1; Length 3747;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 549 ATACTAATAAAAAAAAAAAAAAAAAAAAAA 574
Db 3688 ATACTAATAAAAAAAAAAAAAAAAAAAAAA 3713

RESULT 14
US-09-031-563-1
; Sequence 1, Application US/09031563A
; Patent No. 6022708
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Rozenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: FI272
; CURRENT APPLICATION NUMBER: US/09/031,563A
; CURRENT FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 4880
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unknown N
; LOCATION: 4160, 4243, 4361
; OTHER INFORMATION: unknown
; US-09-031-563-1

Query Match 4.5%; Score 26; DB 3; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAGCGTCCGCCACGCGTCCGC 26
Db 33 ACCCAGCGTCCGCCACGCGTCCGC 58

RESULT 15
US-09-392-277-1
; Sequence 1, Application US/09392277A
; Patent No. 6451977
; GENERAL INFORMATION:

```

```
; APPLICANT: Frederic de Sauvage
; APPLICANT: Airon Rosenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: P1272R1P1
; CURRENT APPLICATION NUMBER: US/09/392,277A
; CURRENT FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US 09/258,000
; EARLIER FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: US 60/076,072
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 28
; SEQ ID NO 1
; LENGTH: 4880
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unknown N
; LOCATION: 4160, 4243, 4361
; OTHER INFORMATION: unknown
US-09-392-277-1

Query Match      4.5%; Score 26; DB 4; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ACCGCGGTCGCCGCCGCGTCCGC 26
        |||||||
Db      33  ACCGCGGTCGCCGCCGCGTCCGC 58

Search completed: June 16, 2003, 08:47:04
Job time : 42.9504 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 03:58:29 ; Search time 188.731 Seconds
(without alignments)
6849.143 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgcgcgcacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapex 60.0 , Gapex 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :	N_Geneseq_101002.*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	31	5.4	908	22	AAF97898 Human secreted pro
2	29	5.1	569	24	ABK62601 Rat sequence diffe
3	29	5.1	1926	22	AAF81796 Human secreted pro
4	28	4.9	296	22	AAK57062 Human immune/haema
5	28	4.9	401	22	AAI87312 Human polynucleoti
6	28	4.9	532	20	AAK22224 Human secreted pro
7	28	4.9	1592	24	ABQ54307 Human polynucleoti
8	28	4.9	2874	22	AAI57944 Human polynucleoti
9	28	4.9	4929	22	AAK51691 Human polynucleoti

10	28	4.9	4929	22	AAI58784 Human polynucleoti
c 11	28	4.9	5928	24	ABL34086 Human immune syste
c 12	28	4.9	6593	24	ABL32478 Chemically treated
c 13	28	4.9	13919	24	ABL22305 Human SNP oligonuc
c 14	27	4.7	51	22	AAI30125 Human colon cancer
15	27	4.7	233	22	AAH33064 Human polynucleoti
16	27	4.7	380	22	AAI93735 Human polynucleoti
17	27	4.7	391	22	AAI88702 Human polynucleoti
18	27	4.7	393	22	AAI88879 Human polynucleoti
19	27	4.7	400	21	AAK98255 Human colon cancer
20	27	4.7	451	22	AAK88891 Human digestive sy
21	27	4.7	451	22	AAI57612 Human colorectal c
22	27	4.7	471	24	ABL67877 Ovary cancer relat
23	27	4.7	474	22	ABA57396 Human foetal liver
24	27	4.7	474	22	ABA26916 Probe #5382 for ge
25	27	4.7	474	22	AAK05432 Human brain expres
26	27	4.7	474	22	AAK31031 Human bone marrow
27	27	4.7	474	22	AAI36941 Probe #5637 used t
28	27	4.7	474	22	ABS05781 Human genome-deriv
c 29	27	4.7	477	22	AAF93500 cDNA encoding SRT
30	27	4.7	487	22	AAI81103 Human polynucleoti
31	27	4.7	492	22	AAI86826 Human polynucleoti
32	27	4.7	531	22	AAK57829 Human polynucleoti
33	27	4.7	534	22	AAI81363 Human immune/haema
34	27	4.7	570	23	ABV59114 Human polynucleoti
c 35	27	4.7	583	24	ABK55196 Human prostate exp
c 36	27	4.7	746	23	AAK56585 Human colon cancer
c 37	27	4.7	937	24	ABQ25590 DNA encoding novel
38	27	4.7	937	24	ABQ25590 Oligonucleotide fo
39	27	4.7	1101	22	AAH34314 Human colon cancer
40	27	4.7	1101	24	ABL90325 Human polynucleoti
41	27	4.7	1263	21	AAK51245 Human TANGO 130 cD
42	27	4.7	1281	22	AAK51610 Human polynucleoti
43	27	4.7	1308	22	AAK08380 Human secreted pro
44	27	4.7	1385	22	AAK84349 Corn clone 7002148
45	27	4.7	1459	24	ABL90182 Human polynucleoti

ALIGNMENTS

RESULT 1	
AAF97898	
ID	AAF97898 standard; cDNA; 908 BP.
XX	
AC	AAF97898;
XX	
DT	01-JUN-2001 (first entry)
XX	
DE	Human secreted protein cDNA, SEQ ID NO: 25.
XX	
KW	Human; secreted protein; immunomodulatory; antisclerotic;
KW	dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
KW	vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW	neotropic; anticonvulsant; antialzheimers; antiparkinsonian;
KW	antimicrobial; vulnery; vaccine; gene therapy; cancer;
KW	protein coordinate data; infection; ss.
OS	Homo sapiens.
XX	
PN	WO200121658-A1.
XX	
PD	29-MAR-2001.
XX	
PF	22-SEP-2000; 2000WO-US26013.
XX	
PR	24-SEP-1999; 99US-0155709.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
PI	Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
PI	Young PE, Wei P, Florence KA;

XX DR WPI; 2001-235311/24.
XX PT Nucleic acids encoding 32 human secreted polypeptides, useful for
XX PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX PT disease and diabetic retinopathy -
XX PS Claim 1; Page 730; 890pp; English.
XX CC The present sequence encodes one of 32 novel human secreted polypeptides.
XX CC The nucleic acid molecules and polypeptides they encode may be used in
XX CC the prevention, diagnosis and treatment of diseases such as
XX CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
XX CC and human immunodeficiency virus (HIV) infections), hyperproliferative
XX CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
XX CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
XX CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
XX CC neovascularisation and diabetic retinopathy), neurological disorders
XX CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
XX CC infectious diseases and/or for promoting wound healing, regeneration
XX CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
XX CC secreted polypeptides. They may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acid
XX CC sequences in samples. The polypeptides may be used as antigens in the
XX CC production of antibodies and in assays to identify modulators of
XX CC their expression and activity.
XX SQ Sequence 908 BP; 242 A; 290 C; 197 G; 175 T; 4 other;
Query Match 5.4%; Score 31; DB 22; Length 908;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCACGCGTCCGCCACGCGTCCGCACGAC 31
Db 19 ACCACGCGTCCGCCACGCGTCCGCACGAC 49
RESULT 2
ABK62601/c
ID ABRK62601 standard; cDNA; 569 BP.
XX AC ABRK62601;
XX CC 18-JUN-2002 (first entry)
XX DT Rat sequence differentially expressed in response to a hepatotoxin #508.
XX DE Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX KW differential expression; centrilobular necrosis; steatosis.
XX OS Rattus norvegicus.
XX PA WO200210453-A2.
XX PD 07-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23872.
XX PR 31-JUL-2000; 2000US-222040P.
XX PR 02-NOV-2000; 2000US-244880P.
XX PR 11-MAY-2001; 2001US-290029P.
XX PR 15-MAY-2001; 2001US-290645P.
XX PR 22-MAY-2001; 2001US-292336P.
XX PR 06-JUN-2001; 2001US-295798P.
XX PR 13-JUN-2001; 2001US-297457P.
XX PR 19-JUN-2001; 2001US-298884P.
XX PR 09-JUL-2001; 2001US-303459P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX PN

DR XX WPI; 2002-241625/29.
PT XX Predicting toxic effects of compounds or the progression of these toxic
PT XX effects by determining the changes in gene expression in tissues or
PT XX cells exposed to the toxin and comparing these to gene expression in
PT XX unexposed tissues or cells -
XX PS Claim 1; Seq ID No 508; 239pp; English.
XX CC The invention relates to methods for predicting toxic effects of
XX CC compounds or the progression of these toxic effects by determining the
XX CC global changes in gene expression in tissues or cells exposed to the
XX CC toxin and comparing these to gene expression in unexposed tissues or
XX CC cells. Also included are methods of predicting at least one toxic
XX CC effect of a compound or progression of a toxic effect, preferably the
XX CC hepatotoxicity of a compound, comprising detecting the level of
XX CC expression in a tissue or cell sample exposed to the compound of two or
XX CC more genes listed in the specification, where differential expression of
XX CC the genes is indicative of at least one toxic effect or progression.
XX CC The method can also be used to identify an agent which modulates the
XX CC toxic response and predict cellular pathways that a compound modulates
XX CC in a cell. The methods utilise a set of at least two probes (on a solid
XX CC support in kit form), where each of the probes comprises a sequence that
XX CC specifically hybridises to a gene listed in the specification, a computer
XX CC system comprising a database containing information identifying the
XX CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX CC set of genes comprising at least two genes listed in the specification,
XX CC and a user interface to view the information used to present information,
XX CC identifying the expression level in a tissue or cell of at least one gene
XX CC listed in the specification. The method is useful for elucidating global
XX CC changes in gene expression and for identifying toxicity markers in
XX CC tissues or cell exposed to a known toxin. The genes may be used as
XX CC toxicity markers in drug screening and toxicity assays. The genes and
XX CC gene expression information may be used as diagnostic markers for the
XX CC prediction or identification of the physiological state of tissue or cell
XX CC sample that has been exposed to a compound or agent. Hepatotoxicity
XX CC is characterised by centrilobular necrosis and steatosis. The present
XX CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX CC which is differentially expressed in response to a hepatotoxic agent.
XX SQ Sequence 569 BP; 132 A; 98 C; 117 G; 222 T; 0 other;
Query Match 5.1%; Score 29; DB 24; Length 569;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 546 GTAATACTAAAAAATAAAAAAATAAAAAA 574
Db 73 GTAATACTAAAAAATAAAAAAATAAAAAA 45
RESULT 3
AAF81796
ID AAF81796 standard; cDNA; 1926 BP.
XX AC AAF81796;
XX CC 12-JUN-2001 (first entry)
XX DT Human secreted protein gene 10 SEQ ID NO:20.
XX DE Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
XX KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
XX KW immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;
XX KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
XX KW antialzheimer; antiparkinsonian; antimicrobial; vulnery; gene therapy;
XX KW immune disorder; hyperproliferative disorder; cardiovascular disease;
XX KW cancer; angiogenic disorder; neurological disorder; infectious disease;
XX KW wound healing; regeneration; chemotaxis; chromosome 20; ss.
XX OS Homo sapiens.
XX PI WO200112775-A2.
XX PN


```

XX PD 22-FEB-2001.
XX PF
XX PR 16-AUG-2000; 2000WO-US22325.
XX PR 17-AUG-1999; 99US-0149182.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PR
XX PI Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
XX PI Birse CE, Young PE, Komatsculis GA, Moore PA, Soppet DR;
XX PI WPI; 2001-147550/15.
XX PR P-PSDB; AAB74742.
XX PR
XX PT Nucleic acids encoding 25 human secreted polypeptides, useful for
XX PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX PT disease and diabetic retinopathy -
XX PR
XX PS Claim 1; Page 446; 485pp; English.
XX CC
XX CC AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
XX CC to AAB74772. Human secreted proteins can have activities based on the
XX CC tissues and cells they are expressed in. Example of activities include:
XX CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
XX CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiac;
XX CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
XX CC anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial; and
XX CC vulnerary. Human secreted proteins can be used in gene therapy and
XX CC vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins
XX CC (PEP1) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. For example, NAM1
XX CC and PEP1 may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patients genome
XX CC that affect the activity of proteins by expressing inactive proteins or
XX CC to supplement the patients own production of polypeptides. Disorders that
XX CC may be prevented, diagnosed and/or treated include immune disorders,
XX CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
XX CC angogenic disorders, neurological disorders, infectious diseases and/or
XX CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
XX CC AAF81786 and AAB74732 represent sequences used in the exemplification of
XX CC the present invention.
XX CC
XX SQ Sequence 1926 BP; 519 A; 429 C; 430 G; 548 T; 0 other;

Query Match 5.1%; Score 29; DB 22; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 574
DB 1881 GTATATCTAAAAAAAAAAAAAAAAAAAAA 1909

RESULT 4
AAK57062
ID AAK57062 standard; cDNA; 296 BP.
XX AC
XX AC AAK57062;
XX DT
XX DT 06-NOV-2001 (first entry)
XX DE
XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2122.
XX KW
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX OS
XX OS Homo sapiens.
XX OS
XX OS WO2001571182-A2.
XX PN
XX PN 09-AUG-2001.
XX PD
XX PD

```

29-SEP-2000; 2000US-02363370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 13-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249222.
 PR 17-NOV-2000; 2000US-0249223.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249229.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 P-PSDB; AAM84281.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX Claim 1; SEQ ID NO 2122; 3071pp + Sequence Listing; English.
 PS
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC the nucleic acids may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells: AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX Sequence 296 BP; 110 A; 66 C; 57 G; 61 T; 2 other;
 SQ

Query Match 4.9%; Score 28; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 0.074; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 0;
 QY 2 CCCACGGTCGCCGCCACGGTCGCCGACA 29
 DB 17 CCCACGGTCGCCGCCACGGTCGCCGACA 44

RESULT 5
 AAI87312
 ID - AAI87312 standard; cDNA; 401 BP.
 XX AAI87312;
 AC AAI87312;
 DT 06-NOV-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 7372.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX Homo sapiens.
 OS WO200164835-A2.
 XX 07-SEP-2001.
 XX 26-FEB-2001; 2001WO-US04927.
 XX 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR P-PSDB; AAO7381.
 XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX Claim 1; SEQ ID NO 7372; 1399pp + Sequence Listing; English.
 PS

XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.

XX Sequence 401 BP; 92 A; 110 C; 78 G; 121 T; 0 other;

Query Match 4.9%; Score 28; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCAGCGTCCGCCACGCGTCCGCACA 29

Db 1 CCCAGCGTCCGCCACGCGTCCGCACA 28

RESULT 6

AA122224

ID AAX22224 standard; DNA; 532 BP.

AC AAX22224;

DT 18-MAY-1999 (first entry)

DE Human secreted protein gene 14 clone HHEPT60.

XX Human; secreted protein; gene therapy; protein therapy; cancer; weight;
 KW tumour; chromosome mapping; forensic; haematological disease; allergy;
 KW inflammation; cell proliferation; viral infection; wound healing;
 KW modulation; appetite; behaviour; food additive; preservative; ss.

OS Homo sapiens.

PN WO9903990-A1.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US14613.

PR 18-AUG-1997; 97US-0056361.

PR 16-JUL-1997; 97US-0052661.

PR 16-JUL-1997; 97US-0052870.

PR 16-JUL-1997; 97US-0052871.

PR 16-JUL-1997; 97US-0052872.

PR 16-JUL-1997; 97US-0052873.

PR 16-JUL-1997; 97US-0052874.

PR 22-JUL-1997; 97US-0052875.

PR 22-JUL-1997; 97US-0053440.

PR 22-JUL-1997; 97US-0053441.

PR 18-AUG-1997; 97US-0053442.

PR 18-AUG-1997; 97US-0055683.

PR 18-AUG-1997; 97US-0055724.

PR 18-AUG-1997; 97US-0055725.

PR 18-AUG-1997; 97US-0055726.

PR 18-AUG-1997; 97US-0055946.

PR 18-AUG-1997; 97US-0055952.

PR 18-AUG-1997; 97US-0055985.

PR 18-AUG-1997; 97US-0055989.

PR 18-AUG-1997; 97US-0056359.

PR (HUMA-) HUMAN GENOME SCI INC.

PA XX

XX

PI

PI Duan R, Feng P, Ferrie AM, Florence KA, Fouad J;

PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE;

XX Yu G;

DR WPI: 1999-132234/11.

DR P-PSDB; AAY01396.

XX

PT New nucleic acids encoding secreted human proteins - potentially
 PT useful for treating and diagnosing diseases and identifying specific
 PT binding agents

XX

PS Claim 4; Page 176; 251pp; English.

XX

XX The invention relates to nucleic acid sequences (AAX22211 to AAX22282)
 CC encoding human secreted proteins (AAY01383 to AAY01454). The secreted
 CC protein gene sequences are deposited with the ATCC under deposit number
 CC ATCC 209138, 209139 or 209141. Host cells containing vectors comprising
 CC the nucleic acid sequences are used for the recombinant expression of
 CC the secreted proteins. The polynucleotide and amino acid sequences are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also diagnosed
 CC by determining the amount of the new polypeptides in a sample or by the
 CC presence of mutations in the new polynucleotides. The nucleic acid
 CC sequences, or its fragments, are useful for chromosome identification
 CC and mapping; as antisense and triplex-forming therapeutics; in gene
 CC therapy; for (forensic) identification of individuals; as molecular
 CC weight markers; to identify related sequences or specific mRNA; in
 CC preparation of oligomers and to raise anti-DNA antibodies. Antibodies are
 CC useful as immunoassay reagents (including for in vivo imaging) and
 CC therapeutically to inhibit or activate particular polypeptides. A very
 CC wide range of disorders may be treated with the polynucleotide and
 CC polypeptide sequences, e.g. autoimmune or haematological diseases,
 CC allergy, inflammation, cancer or other forms of cell proliferation, viral
 CC or other infections. The sequences may also be useful in wound healing,
 CC to modulate differentiation of embryonic stem cells, to modulate weight,
 CC appetite, behaviour etc. and as food additive or preservative. The
 CC present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).

SQ Sequence 532 BP; 113 A; 146 C; 126 G; 147 T; 0 other;

Query Match 4.9%; Score 28; DB 20; Length 532;

Best Local Similarity 100.0%; Pred. No. 0.066;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGCAC 28

Db 51 ACCACGCGTCCGCCACGCGTCCGCAC 78

RESULT 7

ABQ54307

ID ABQ54307 standard; cDNA; 1592 BP.

XX ABQ54307;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HPSI74 cDNA, SEQ ID NO:187.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 3q21-25;
 KW gene; ss.

XX Homo sapiens.

OS

XX	WO200200677-A1.	Human polynucleotide SEQ ID NO 147.
XX	03-JAN-2002.	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
XX	07-JUN-2001; 2001WO-US18569.	Homo sapiens.
XX	07-JUN-2000; 2000US-209467P.	WO200153312-A1.
XX	(HUMA-) HUMAN GENOME SCI INC.	26-JUL-2001.
XX	Birse CE, Rosen CA;	26-DEC-2000; 2000WO-US34263.
XX	WPI: 2002-147878/19.	21-JAN-2000; 2000US-0488725.
XX	P-PSDB; ABP41230.	25-APR-2000; 2000US-052317.
XX	Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases -	09-JUL-2000; 2000US-0598042.
XX	Claim 1: SEQ ID No 187; 2922pp; English.	19-JUL-2000; 2000US-0620312.
XX	The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP4328) and to cDNAs encoding them (ARQ54131-ARQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents cDNA encoding a human ovarian antigen of the invention.	03-AUG-2000; 2000US-0653450.
XX	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.	14-SEP-2000; 2000US-0862191.
XX	Sequence 1592 BP; 458 A; 322 C; 338 G; 474 T; 0 other;	19-OCT-2000; 2000US-0693036.
XX	Query Match 4.9%; Score 28; DB 24; Length 1592;	29-NOV-2000; 2000US-0727344.
XX	Best Local Similarity 100.0%; Pred. No. 0.054;	(HYSE-) HYSEQ INC.
XX	Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX	1 ACCCACGGTCCGCCACGGTCCGCAC 28	PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX	4 ACCCACGGTCCGCCACGGTCCGCAC 31	PI Zhao QJ, Zhou P, Goodrich R, Drmanac RT;
XX	RESULT 8	XX WPI: 2001-442253/47.
XX	AAI57944	XX P-PSDB; AAM38788.
XX	AAI57944 standard; cDNA: 2674 BP.	XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
XX	AAI57944;	XX Claim 1: SEQ ID NO 147; 10078pp; English.
XX	22-OCT-2001 (first entry)	XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAI62213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.
XX	Sequence 1592 BP; 458 A; 322 C; 338 G; 474 T; 0 other;	XX C.N.S disorders.
XX	Query Match 4.9%; Score 28; DB 24; Length 1592;	XX Note: The sequence data for this patent did not form part of the printed specification.
XX	Best Local Similarity 100.0%; Pred. No. 0.054;	XX Sequence 2674 BP; 906 A; 529 C; 580 G; 659 T; 0 other;
XX	Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX Query Match 4.9%; Score 28; DB 22; Length 2674;
XX	1 ACCCACGGTCCGCCACGGTCCGCAC 28	XX Best Local Similarity 100.0%; Pred. No. 0.049;
XX	4 ACCCACGGTCCGCCACGGTCCGCAC 31	XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	RESULT 8	XX 1 ACCCACGGTCCGCCACGGTCCGCAC 28
XX	AAI57944	XX 21 ACCCACGGTCCGCCACGGTCCGCAC 48
XX	AAI57944 standard; cDNA: 2674 BP.	XX RESULT 9
XX	AAI57944;	XX AAK51691
XX	22-OCT-2001 (first entry)	XX ID AAK51691 standard; cDNA: 4929 BP.
XX		XX AAK51691;
XX		XX AC AAK51691;

Human polynucleotide SEQ ID NO 147.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-052317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QJ, Zhou P, Goodrich R, Drmanac RT;

WPI: 2001-442253/47.
P-PSDB; AAM38788.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

Claim 1: SEQ ID NO 147; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 2674 BP; 906 A; 529 C; 580 G; 659 T; 0 other;

Query Match 4.9%; Score 28; DB 22; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGGTCGCGCAC 28
| | | | | | | | | | | | | | |
Db 21 ACCACGCGTCCGCCACGGTCGCGCAC 48

RESULT 9
ID AAI57944 standard; cDNA: 2674 BP.
XX AAI57944;
XX 22-OCT-2001 (first entry)
AC

XX DT 06-NOV-2001 (first entry)
 XX DE Human polynucleotide SEQ ID NO 236.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX OS Homo sapiens.
 XX PN WO200157190-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US04098.
 XX PR 03-FEB-2000; 2000US-0496914.
 XX PR 27-APR-2000; 2000US-0560875.
 XX PR 20-JUN-2000; 2000US-0598075.
 XX PR 19-JUL-2000; 2000US-0620325.
 XX PR 01-SEP-2000; 2000US-0654936.
 XX PR 15-SEP-2000; 2000US-0663561.
 XX PR 20-OCT-2000; 2000US-0693325.
 XX PR 30-NOV-2000; 2000US-0728422.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM78558.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS Claim 1; Page 1115-1120; 6221pp; English.
 XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX SQ Sequence 4929 BP; 1676 A; 904 C; 981 G; 1368 T; 0 other;
 Query Match 4.9%; Score 28; DB 22; Length 4929;
 Best Local Similarity 100.0%; Pred. No. 0.044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCACGCGTCCGCCACGCGTCCGCAC 28
 Db ||||||||||||||||||||||||||||
 45 ACCACGCGTCCGCCACGCGTCCGCAC 72
 RESULT 10
 AAI58784
 ID AAI58784 standard; cDNA; 4929 BP.
 XX AC AAI58784;
 XX

DT 22-OCT-2001 (first entry)
 XX DE Human polynucleotide SEQ ID NO 987.
 XX KW Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia; ss.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US34263.
 XX PR 21-JAN-2000; 2000US-0488725.
 XX PR 25-APR-2000; 2000US-0552317.
 XX PR 09-JUL-2000; 2000US-0598042.
 XX PR 19-JUL-2000; 2000US-0620312.
 XX PR 03-AUG-2000; 2000US-0653450.
 XX PR 14-SEP-2000; 2000US-0662191.
 XX PR 19-OCT-2000; 2000US-0693036.
 XX PR 29-NOV-2000; 2000US-0727344.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 DR WPI: 2001-442253/47.
 DR P-PSDB; AAM39628.
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 987; 10078pp; English.
 XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM39642-AAM42213) with nototropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral neuropathies and central nervous system diseases,
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX SQ Sequence 4929 BP; 1678 A; 902 C; 982 G; 1367 T; 0 other;
 Query Match 4.9%; Score 28; DB 22; Length 4929;
 Best Local Similarity 100.0%; Pred. No. 0.044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCACGCGTCCGCCACGCGTCCGCAC 28
 Db ||||||||||||||||||||||||||||
 45 ACCACGCGTCCGCCACGCGTCCGCAC 72
 RESULT 11
 ABL34086/C
 ID ABL34086 standard; DNA; 5928 BP.

XX ABL34086;
 XX 26-MAR-2002 (first entry)
 XX Human immune system associated gene SEQ ID NO: 2059.
 XX Human: immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 XX for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation
 XX Claim 1; SEQ ID NO 2059; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 XX can be used in the diagnosis and treatment of immune system disorders,
 XX including eye diseases such as retinopathy, neovascular glaucoma and
 XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 XX diseases. The present sequence is a gene of the invention.
 XX Sequence 5928 BP; 1649 A; 110 C; 1223 G; 2946 T; 0 other;
 SQ
 Query Match 4.9%; Score 28; DB 24; Length 5928;
 Best Local Similarity 100.0%; Pred. No. 0.043;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 574
 DB 5865 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 5838
 RESULT 12
 ABL32478/c
 ID ABL32478 standard; DNA; 6593 BP.
 XX ABL32478;
 AC
 XX 26-MAR-2002 (first entry)
 DT Human immune system associated gene SEQ ID NO: 451.
 XX Human: immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 XX for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation
 XX Claim 1; SEQ ID NO 2059; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 XX can be used in the diagnosis and treatment of immune system disorders,
 XX including eye diseases such as retinopathy, neovascular glaucoma and
 XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 XX diseases. The present sequence is a gene of the invention.
 XX Sequence 5928 BP; 1649 A; 110 C; 1223 G; 2946 T; 0 other;
 SQ
 Query Match 4.9%; Score 28; DB 24; Length 5928;
 Best Local Similarity 100.0%; Pred. No. 0.043;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 574
 DB 5865 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 5838
 RESULT 13
 ABL92305/c
 ID ABL92305 standard; DNA; 13919 BP.
 XX ABL92305;
 AC
 XX 01-JUL-2002 (first entry)
 DT Chemically treated DNA repair gene fragment complementary to#57.
 XX DNA repair; cytosine methylation; PMS2L1; PMS2L2; PMS2L3;
 KW PMS2; L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPP1; RFC4;
 KW DDT1L; FANCB; XRCC8; ataxia telangiectasia; aging; Bloom's syndrome;
 KW Cockayne syndrome; Nijmegen breakage syndrome; Werner syndrome;
 KW immunodeficiency; trichothiodystrophy; Fanconi's anaemia; solid tumour;
 KW cancer; ds.
 XX Unidentified.
 OS
 XX WO200181622-A2.
 PN
 XX 01-NOV-2001.
 PD
 XX 06-APR-2001; 2001WO-EP03972.
 XX

KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 XX for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation
 XX Claim 1; SEQ ID NO 451; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 XX can be used in the diagnosis and treatment of immune system disorders,
 XX including eye diseases such as retinopathy, neovascular glaucoma and
 XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 XX diseases. The present sequence is a gene of the invention.
 XX Sequence 6593 BP; 1732 A; 121 C; 1448 G; 3292 T; 0 other;
 SQ
 Query Match 4.9%; Score 28; DB 24; Length 6593;
 Best Local Similarity 100.0%; Pred. No. 0.042;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 574
 DB 1002 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 975
 RESULT 13
 ABL92305/c
 ID ABL92305 standard; DNA; 13919 BP.
 XX ABL92305;
 AC
 XX 01-JUL-2002 (first entry)
 DT Chemically treated DNA repair gene fragment complementary to#57.
 XX DNA repair; cytosine methylation; PMS2L1; PMS2L2; PMS2L3;
 KW PMS2; L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPP1; RFC4;
 KW DDT1L; FANCB; XRCC8; ataxia telangiectasia; aging; Bloom's syndrome;
 KW Cockayne syndrome; Nijmegen breakage syndrome; Werner syndrome;
 KW immunodeficiency; trichothiodystrophy; Fanconi's anaemia; solid tumour;
 KW cancer; ds.
 XX Unidentified.
 OS
 XX WO200181622-A2.
 PN
 XX 01-NOV-2001.
 PD
 XX 06-APR-2001; 2001WO-EP03972.
 XX

PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPITG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-034446/04.
 DR WPI; 2002-034446/04.
 XX New nucleic acid derived from genes associated with DNA repair, useful
 PT for diagnosis, e.g. of ataxia telangiectasia, by determination of
 PT cytosine methylation.
 XX Claim 1; SEQ ID NO 114; 25pp + sequence listing; English.
 PS
 CC The invention relates to nucleic acids containing a sequence of at least
 CC 18 nucleotides of chemically treated DNA of genes associated with DNA
 CC repair, and their complements. The invention also relates to nucleic
 CC acids comprising at least 18 base pairs of the chemically pretreated DNA
 CC of genes associated with DNA repair selected from PMS2L1, PMS2L12,
 CC PMS2L2, PMS2L3, PMS2, L4, PMS2L5, PMS2L6, MGMT, MSH2, NUDT1, TDG, INPPL1,
 CC RFC4, DDIRIL, FANCB, or XRCC8. Nucleic acids of the invention and related
 CC oligomers, are useful for diagnosis of diseases associated with gene
 CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
 CC Cockayne syndrome, Mijmegen breakage syndrome or Werner's syndrome,
 CC immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours
 CC and cancer, particularly by determining status of cytosine methylation
 CC and/or by detecting single-nucleotide polymorphisms. Determination of
 CC individual methylation patterns may allow development of individualised
 CC therapies. The sequences given in records ABL92192-ABL92335 represent
 CC chemically pre-treated DNA fragments from genes associated with DNA
 CC repair, and their complements.
 CC Note: The sequence data for this patent is not represented in the
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 13919 BP; 3717 A; 296 C; 3349 G; 6557 T; 0 other;
 Query Match 4.9%; Score 28; DB 24; Length 13919;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTAAAAA...AAAAAAAAAAAAA 574
 Db 7862 TAATCTAAAAA...AAAAAAAAAAAAA 7835
 RESULT 14
 AAL30125/c
 ID AAL30125 standard; DNA; 51 BP.
 XX
 AC AAL30125;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human SNP oligonucleotide #3333.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW anyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinein; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.
 XX 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX (CURA-) CURAGEN CORP.
 PA Shimkets RA, Leach M;
 XX WPI; 2001-465210/50.
 DR
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.
 XX Claim 1; Page 2342; 4143pp; English.
 PS
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kineins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, cancer
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 XX
 SQ Sequence 51 BP; 11 A; 1 C; 3 G; 36 T; 0 other;
 Query Match 4.7%; Score 27; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 548 AATCTAAAAA...AAAAAAAAAAAAA 574
 Db 32 AATCTAAAAA...AAAAAAAAAAAAA 6
 RESULT 15
 AAH33064
 ID AAH33064 standard; cDNA; 233 BP.
 XX
 AC AAH33064;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:120.
 XX
 DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; ss.
 KW
 KW Homo sapiens.
 OS
 XX WO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US26524.
 PF
 XX 29-SEP-1999; 99US-0157137.
 PR
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI

XX WPI: 2001-235357/24.
DR P-PSDB; AG73633.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 1; Page 2294; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AG73514 to AG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 233 BP; 116 A; 37 C; 40 G; 40 T; 0 other;

Query Match 4.7%; Score 27; DB 22; Length 233;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACCGCTCCGCA 27
Db |||||||||||||||||||||||||
4 ACCACGCGTCCGCCACCGCTCCGCA 30

Search completed: June 16, 2003, 06:54:05
Job time : 192.231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 07:57:31 ; Search time 127.198 Seconds
(without alignments)
6534.570 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 acccagcgctccgcccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 1029858 seqs, 724030393 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	574	100.0	574	9	US-10-090-035-3
2	574	100.0	577	9	US-10-090-035-5
3	313	54.5	676	9	US-10-090-035-1
4	194	33.8	529	9	US-10-090-035-9
5	178	31.0	580	9	US-10-090-035-7
6	139	24.2	524	9	US-10-090-035-17
7	91	15.9	237	10	US-09-923-876-2788
8	44	7.7	348	9	US-10-090-035-13
9	44	7.7	591	9	US-10-090-035-15
10	34	5.9	436	9	US-10-090-035-21
11	34	5.9	584	9	US-10-090-035-19
12	31	5.4	584	9	US-10-090-035-23
13	29	5.1	538	10	US-09-800-729-25
14	29	5.1	569	10	US-09-918-995-17655
15	28	4.9	405	9	US-09-917-800A-508
16	28	4.9	481	9	US-09-918-995-16259
17	28	4.9	532	9	US-09-918-995-6561
18	28	4.9	532	9	US-09-776-724A-24
19	28	4.9	816	9	US-10-125-237-28

20	28	4.9	816	9	US-10-105-891-28
21	28	4.9	2674	9	US-10-098-841-148
22	28	4.9	4929	9	US-10-037-270-674
23	27	4.7	390	9	US-09-918-995-17194
24	27	4.7	393	9	US-09-918-995-7993
25	27	4.7	394	9	US-10-060-036-2561
26	27	4.7	400	9	US-09-925-239-465
27	27	4.7	400	10	US-09-925-239-265
28	27	4.7	421	9	US-09-918-995-16314
29	27	4.7	421	9	US-09-918-995-2984
30	27	4.7	451	9	US-10-072-349-76
31	27	4.7	451	10	US-09-764-855-76
32	27	4.7	468	9	US-09-967-768A-69
33	27	4.7	471	10	US-09-864-761-5382
34	27	4.7	474	10	US-09-918-995-16810
35	27	4.7	488	9	US-09-918-995-16651
36	27	4.7	508	9	US-09-918-995-16651
37	27	4.7	583	10	US-09-919-580-666
38	27	4.7	1263	10	US-09-919-580-666
39	27	4.7	1794	10	US-09-925-301-426
40	27	4.7	1837	9	US-09-813-153-72
41	27	4.7	1927	9	US-10-138-846-11355
42	27	4.7	7802	9	US-10-239-676-56
43	27	4.7	10445	9	US-09-764-891-6380
44	27	4.7	16892	9	US-09-764-872-642
45	26	4.5	164	10	US-09-925-302-27

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090, 035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Query Match 100.0%; Score 574; DB 9; Length 574;
Best Local Similarity 100.0%; Pred. No. 2.5e-278;
Matches 574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAGCGTCCGCCAGCGCTCCGACAGCAATCCACCAAGCACTTCGACGTCACACG 60
DB 1 ACCCAGCGTCCGCCAGCGCTCCGACAGCAATCCACCAAGCACTTCGACGTCACACG 60
QY 61 GCGCTGCGCAGCAGACACACACGCGTCCGACGCAATCCACCAAGCACTTCGACGTCACACG 120
DB 61 GCGCTGCGCAGCAGACACACACGCGTCCGACGCAATCCACCAAGCACTTCGACGTCACACG 120
QY 121 ACTGCTCGAGGAGGTGAGGTGCGTGGCCCGCGCGCTTCGCGCCGCGCGCGCGCGCGG 180
DB 121 ACTGCTCGAGGAGGTGAGGTGCGTGGCCCGCGCGCTTCGCGCCGCGCGCGCGCGCGG 180
QY 181 TCCAGCAGCAGTCTGCTCAAGGAGAGTTCGAGGAGGTTCGACAGGCTCTACGCGCGCGCGG 240
DB 181 TCCAGCAGCAGTCTGCTCAAGGAGAGTTCGAGGAGGTTCGACAGGCTCTACGCGCGCGCGG 240


```
Db 31 CACAGACACACCAAGCGTGGCGCCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 90
QY 130 AGGAGGTGAGGTGCGTGGCGCCCGCGGCTTCGGCGCCGACAGCGCGCGGCGTCCAGCAGC 189
Db 91 AGGAGGTGAGGTGCGTGGCGCCCGCGGCTTCGGCGCCGACAGGAGCGCGTCCAGCAGC 150
QY 190 AGTGTCTAAGAGAGTTCGAGGAGTTCGACACGCTCTCAGCGCGCGCGCAACACACC 249
Db 151 AGTGTCTAAGAGAGTTCGAGGAGTTCGACACGCTCTCAGCGCGCGCGCAACACACC 210
QY 250 ACCACCAACCA 259
Db 211 ACCACCAACCA 220

RESULT 7
US-09-923-876-2788
; Sequence 2788, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2788
; LENGTH: 237
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: incyte ID No. US20020013958A1 700161180H1
; LOCATION: 125-146, 202
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-2788

Query Match 15.9%; Score 91; DB 10; Length 237;
Best Local Similarity 100.0%; Pred. No. 6.2e-36;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 CACAGACACACCAAGCGTGGCGCCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 129
Db 34 CACAGACACACCAAGCGTGGCGCCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 93
QY 130 AGGAGGTGAGGTGCGTGGCGCCCGCGGCTTCGGCGCCGACAGCGCGCGTCCAGCAGC 160
Db 94 AGGAGGTGAGGTGCGTGGCGCCCGCGGCTTCGGCGCCGACAGCGCGCGTTCAGCAGC 124

RESULT 8
US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc.feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-13

Query Match 7.7%; Score 44; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 146
Db 59 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 102

RESULT 9
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc.feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15

Query Match 7.7%; Score 44; DB 9; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 146
Db 68 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 111

RESULT 10
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
```

```
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc.feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-13

Query Match 7.7%; Score 44; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 146
Db 59 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 102

RESULT 9
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc.feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15

Query Match 7.7%; Score 44; DB 9; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 146
Db 68 ACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTG 111

RESULT 10
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
```

```
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(326)
US-10-090-035-21

Query Match
Best Local Similarity 5.9%; Score 34; DB 9; Length 436;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 140
Db 65 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 98

RESULT 11
US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19

Query Match
Best Local Similarity 5.9%; Score 34; DB 9; Length 584;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 140
Db 57 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 90

RESULT 12
US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19

Query Match
Best Local Similarity 5.9%; Score 34; DB 9; Length 584;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 140
Db 57 CCAGGAGGTGACTCTCTCGGAGGAGGTGAGG 90

RESULT 13
US-09-800-729-25
; Sequence 25, Application US/09800729.
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (891)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (896)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-800-729-25

Query Match
Best Local Similarity 5.4%; Score 31; DB 10; Length 908;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAGCGTCCGCCACGCGTCCGCACAGC 31
Db 19 ACCCAGCGTCCGCCACGCGTCCGCACAGC 49

RESULT 14
US-09-918-995-17655/c
; Sequence 17655, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17655
```

LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(538)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17655

Query Match 5.1%; Score 29; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAGCGTCGCCCGCAGCGTCGCCGACA 29
Db 209 ACCCAGCGTCGCCCGCAGCGTCGCCGACA 181

RESULT 15

US-09-917-800A-508/c
Sequence 508, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 508
LENGTH: 569
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 AI007824
US-09-917-800A-508

Query Match 5.1%; Score 29; DB 10; Length 569;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTAATACTAAAAA...AAAAA 574
Db 73 GTAATACTAAAAA...AAAAA 45

Search completed: June 16, 2003, 10:24:26
Job time : 129.198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 05:14:00 ; Search time 1370.39 Seconds
(Without alignments)
7989.092 Million cell updates/sec

Title: US-10-090-035-1

Perfect score: 676

Sequence: 1 accacgcgtccgcacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_esti:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	308	45.6	500	10	AW787732
2	305	45.1	524	10	AW288876
3	279	41.3	465	10	BE129897
4	268	39.6	523	10	AW331212
5	261	38.6	553	11	AX104409
6	261	38.6	648	14	BQ619167

7	261	38.6	648	14	BQ619315
8	261	38.6	648	14	BQ619318
9	261	38.6	648	14	BQ619337
10	261	38.6	648	14	BQ619383
11	261	38.6	648	14	BQ619390
12	256	37.9	550	12	BG840383
13	255	37.7	546	9	AA979839
14	237	35.1	417	13	BM501439
15	232	34.3	436	10	AW787315
16	226	33.4	552	10	BE025302
17	214	31.7	466	9	AI964534
18	214	31.7	496	10	AW787314
19	200	29.6	539	9	AI855425
20	198	29.3	481	9	AI964458
21	197	29.1	443	10	BE129644
22	188	27.8	401	10	AW289056
23	168	24.9	476	10	BE519299
24	152	22.5	326	10	BE025303
25	137	20.3	274	10	AW288875
26	122	18.0	173	14	BM885464
27	122	18.0	215	13	BM530859
28	122	18.0	220	13	BM154814
29	122	18.0	222	14	BQ085944
30	122	18.0	244	13	BM154314
31	122	18.0	1011	14	BQ954179
32	122	18.0	1030	14	BQ222049
33	121	17.9	240	13	BJ389253
34	121	17.9	242	10	AW781584
35	121	17.9	242	10	BE059705
36	121	17.9	275	12	BE777697
37	121	17.9	305	12	BE609663
38	121	17.9	384	9	AI226246
39	121	17.9	399	9	AI226246
40	121	17.9	402	12	BE726297
41	121	17.9	416	10	AW885721
42	121	17.9	711	12	BG109125
43	121	17.9	768	10	AV755678
44	121	17.9	806	10	AV757705
45	121	17.9	842	10	AV757327

ALIGNMENTS

RESULT 1
AW787732
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AW787732
945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
(SK) zea mays cDNA, mRNA sequence.
AW787732
AW787732.1 GI:7844510
EST.
ze mays.
ze mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 500)
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1..500
/organism="Zea mays"

500 bp mRNA linear EST 16-MAY-2000

16-MAY-2000

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES
source
1. 524
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site: 1; EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

BASE COUNT 107 a 170 c 157 g 89 t 1 others

Query Match 45.1%; Score 305; DB 10; Length 524;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACAGGAGTGGACTACTGCTCGGAGGAGTGGTGGGCC 142
DB 67 GCACCAATGGCTTACTACAGGAGTGGACTACTGCTCGGAGGAGTGGTGGGCC 126
QY 143 CCGGCCGGCTTCGGCCGCCACGGCGGCGTCCAGCAGCAGCTGCTCAAGGAGAGTTC 202
DB 127 CCGGCCGGCTTCGGCCGCCACGGCGGCGTCCAGCAGCAGCTGCTCAAGGAGAGTTC 186
QY 203 GAGGAGTTCGACACGGTATCACCGCGGCGGCCAACCCACCACCATGTTCAACAC 262
DB 187 GAGGAGTTCGACACGGTATCACCGCGGCGGCCAACCCACCACCATGTTCAACAC 246
QY 263 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 322
DB 247 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 306
QY 323 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 382
DB 307 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 366
QY 383 CCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 442
DB 367 CCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 426
QY 443 TTGTGAGGCTATCGTACTTGGCTATCGTACGTGCGACGCTAGCTCCTGTAGCAATT 502
DB 427 TTGTGAGGCTATCGTACTTGGCTATCGTACGTGCGACGCTAGCTCCTGTAGCAATT 486
QY 503 ACACAATAAGCTCGTGGTACCTGAATAAACTTCTTCTGT 540
DB 487 ACACAATAAGCTCGTGGTACCTGAATAAACTTCTTCTGT 524

RESULT 3
BE129897
LOCUS
DEFINITION 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897
VERSION BE129897.1 GI:8577260
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site: 1; EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

BASE COUNT 97 a 159 c 156 g 88 t

Query Match 45.6%; Score 308; DB 10; Length 500;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 456; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACAGGAGTGGACTACTGCTCGGAGGAGTGGTGGGCC 142
DB 29 GCACCAATGGCTTACTACAGGAGTGGACTACTGCTCGGAGGAGTGGTGGGCC 88
QY 143 CCGGCCGGCTTCGGCCGCCACGGCGGCGTCCAGCAGCAGCTGCTCAAGGAGAGTTC 202
DB 89 CCGGCCGGCTTCGGCCGCCACGGCGGCGTCCAGCAGCAGCTGCTCAAGGAGAGTTC 148
QY 203 GAGGAGTTCGACACGGTATCACCGCGGCGGCCAACCCACCACCATGTTCAACAC 262
DB 149 GAGGAGTTCGACACGGTATCACCGCGGCGGCCAACCCACCACCATGTTCAACAC 208
QY 263 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 322
DB 209 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 268
QY 323 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 382
DB 269 GCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 328
QY 383 CCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 442
DB 329 CCGGCATCCACGGCTTCGTTGGTGGCGGAGAGCTTCCTCGCCAGGCTTAAGTGGTGG 388
QY 443 TTGTGAGGCTATCGTACTTGGCTATCGTACGTGCGACGCTAGCTCCTGTAGCAATT 502
DB 389 TTGTGAGGCTATCGTACTTGGCTATCGTACGTGCGACGCTAGCTCCTGTAGCAATT 448
QY 503 ACACAATAAGCTCGTGGTACCTGAATAAACTTCTTCTGT 543
DB 449 ACACAATAAGCTCGTGGTACCTGAATAAACTTCTTCTGT 489

RESULT 2
AW288876
LOCUS
DEFINITION 524 bp mRNA linear EST 16-JAN-2000
707009E07.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW288876
VERSION AW288876
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 524)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences


```
clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 465)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1..465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/notes="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 88 a 151 c 149 g 77 t
ORIGIN

Query Match 41.3%; Score 279; DB 10; Length 465;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
Db 34 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 93

QY 143 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 202
Db 94 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 153

QY 203 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 262
Db 154 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 213

QY 263 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 322
Db 214 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 273

QY 323 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 382
Db 274 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 333

QY 383 CCGCATCCAGCGCGCTTCGTGGTGGCGGTTCGTATGTATGTCTGTGTTGACTGG 442
Db 334 CCGCATCCAGCGCGCTTCGTGGTGGCGGTTCGTATGTATGTCTGTGTTGACTGG 393

QY 443 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 502
Db 394 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 502

QY 503 ACCACAAATAGC 514
Db 454 ACCACAAATAGC 465

RESULT 4
AW331212
LOCUS

DEFINITION
1 (bases 1 to 465)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1..465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/notes="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 88 a 151 c 149 g 77 t
ORIGIN

Query Match 41.3%; Score 279; DB 10; Length 465;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
Db 34 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 93

QY 143 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 202
Db 94 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 153

QY 203 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 262
Db 154 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 213

QY 263 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 322
Db 214 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 273

QY 323 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 382
Db 274 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 333

QY 383 CCGCATCCAGCGCGCTTCGTGGTGGCGGTTCGTATGTATGTCTGTGTTGACTGG 442
Db 334 CCGCATCCAGCGCGCTTCGTGGTGGCGGTTCGTATGTATGTCTGTGTTGACTGG 393

QY 443 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 502
Db 394 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 502

QY 503 ACCACAAATAGC 514
Db 454 ACCACAAATAGC 465

RESULT 4
AW331212
LOCUS

707049E04.x1.707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
AW331212
707049E04.x1.707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
AW331212.1 GI:6827569
EST.
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot, V.
1 (bases 1 to 523)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707049 row: E column: 04.
Location/Qualifiers
1..523
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_host="DH10B"
/notes="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT 125 a 158 c 149 g 91 t
ORIGIN

Query Match 39.6%; Score 268; DB 10; Length 523;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 468; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
Db 29 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 88

QY 143 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 202
Db 89 CCGGCCGGCTTCGGCGCCAGCGCGGGTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 148

QY 203 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 262
Db 149 GAGGAGTTCGACAGGTATACCGCGCGCGGCCAACCCACCACCATGCTGTCACCC 208

QY 263 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 322
Db 209 GCGGCCACAGCTTCGTGGTGGCGAGACCGAGGTTCGAGGAGGACATCAACCTGCACC 268

QY 323 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 382
Db 269 GGCAGGTTCACAGCGCAGGAGAGTTCCTCGCCAGGCTTAACGTAGCGCGCGCGG 328

QY 383 CCGCATCCAGCGCGCTTCGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
Db 329 CCGCATCCAGCGCGCTTCGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 388

QY 443 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 502
Db 389 TTGTGAGGTCATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCTATCTGCT 448
```

QY	503	ACGACATAAGTCGTGGACCTGAATAAAACTTCTTCGTAATACTAATACTA 554
DB	449	ACGACATAAGTCGTGGACCTGAATAAAACTTCTTCGTAATACTAATACTA 500
RESULT 5		
LOCUS	AY104409	553 bp mRNA linear HTC 25-MAY-2002
DEFINITION	Zea mays PC0124784 mRNA sequence.	
ACCESSION	AY104409	
VERSION	AY104409.1	GI:21207487
KEYWORDS	HTC.	
SOURCE	Zea mays.	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC	
AUTHORS	Clade; Panicoideae; Andropogoneae; Zea.	
TITLE	1 (bases 1 to 553)	
JOURNAL	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,	
AUTHORS	Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.	
JOURNAL	Maize Mapping Project/DuPont Consensus Sequences for Design of	
FEATURES	Overgo Probes	
source	2 (bases 1 to 553)	
	Coe, E.C.	
	Direct Submission	
	Submitted (25-APR-2002) Maize Mapping Project, University of	
	Missouri, Columbia, MO 65211, USA	
	Location/Qualifiers	
	1. 553	
	/organism="Zea mays"	
	/db_xref="taxon:637271"	
	/db_xref="taxon:4577"	
	/clone="PC0124784"	
	/clone_lib="Maize Mapping Project/DuPont Consensus	
	Library"	
	/note="this sequence is part of a project of EST	
	assemblies resulting from the application of public	
	contigs to seed DuPont contigs; this resource was	
	assembled by DuPont as part of a collaboration for the	
	overgo addressing of BACs in conjunction with the Maize	
	Mapping Project"	
BASE COUNT	128 a 173 c 162 g 90 t.	
ORIGIN		
Query Match	38.6%;	Score 261; DB 11; Length 553;
Best Local Similarity	99.7%;	Pred. No. 0;
Matches 311;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
QY	237	ACCACACACACCATGGTACACACGGCGGCACGGCTTCGTGGTGGCGAGACCGAGG 296
DB	227	ACCACACACACCATGGTACACACGGCGGCACGGCTTCGTGGTGGCGAGACCGAGG 286
QY	297	TCGAGGAGACATCAACACCTCCACCGCGAGGTCACGAGCGGAGAGAGCTTCCTCG 356
DB	287	TCGAGGAGACATCAACACCTCCACCGCGAGGTCACGAGCGGAGAGAGCTTCCTCG 346
QY	357	CCAGGCTTAACAGAGCGCGCGCGGCATCCACGCCGCTTCGTGGTGGCGCGTG 416
DB	347	CCAGGCTTAACAGAGCGCGCGCGGCATCCACGCCGCTTCGTGGTGGCGCGTG 406
QY	417	CCTTATGTATGTCTGTGGTTGACTGGTTGTGAGGTCATCGTACTTGGCTATCGTAGT 476
DB	407	CCTTATGTATGTCTGTGGTTGACTGGTTGTGAGGTCATCGTACTTGGCTATCGTAGT 466
QY	477	GCAGGCATCAGCTCCTGTACGAATATACGACAAATAGCTCGTGAGCTGAATAAACTTCT 536
DB	467	GCAGGCATCAGCTCCTGTACGAATATACGACAAATAGCTCGTGAGCTGAATAAACTTCT 526
QY	537	TCGTAATACTAA 548
DB	527	TCGTAATACTAA 538

```
ACCESSION BQ619315
VERSION BQ619315.1 GI:21621309
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
REFERENCE 1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

FEATURES
source
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6D01_SK.abl"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN
Query Match 38.6%; Score 261; DB 14; Length 648;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 237 ACCACCACACCATGCTCACCACCGCGCCAGCGCTTCGTGGTCCGCGAGACGAGG 296
Db 233 ACCACCACACCATGCTCACCACCGCGCCAGCGCTTCGTGGTCCGCGAGACGAGG 292
QY 297 TCGAGGAGGACATCAACACCTGCACCGCGCGAGGTCCACGAGCGGAGAGCTTCCTCG 356
Db 293 TCGAGGAGGACATCAACACCTGCACCGCGCGAGGTCCACGAGCGGAGAGCTTCCTCG 352
QY 357 CCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 416
Db 353 CCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 412
QY 417 CCTTATGATGCTGTGTTGACTGTTGTGTCAGGGTCATCGTACTTGGCTATCGTACGT 476
Db 413 CCTTATGATGCTGTGTTGACTGTTGTGTCAGGGTCATCGTACTTGGCTATCGTACGT 472
QY 477 GCAGGCACTCAGCTCCTGTACGAATACGACAATAAGCTCGTACCTGAATAAACTTCT 536
Db 473 GCAGGCACTCAGCTCCTGTACGAATACGACAATAAGCTCGTACCTGAATAAACTTCT 532
QY 537 TCGTAATACTAA 548
Db 533 TCGTAATACTAA 544

RESULT 9
BQ619337
LOCUS BQ619337
DEFINITION RNOSEQ6E12_SK.abl Salt stressed Zea mays roots cDNA library Zea
mays cDNA clone RNOSEQ6E12_SK.abl similar to No homology, mRNA
sequence.
ACCESSION BQ619337
VERSION BQ619337.1 GI:21621331
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
REFERENCE 1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology

ACCESSION BQ619318
VERSION BQ619318.1 GI:21621312
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
REFERENCE 1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
```


Mon Jun 16 14:55:52 2003

source
1. .546
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="WEST2-B7"
/clone_lib="ISUM2"
/tissue_type="above ground tissues"
/lab_host="XLI-MFR Blue"
/note="organ: green seedlings; Vector: pAD-GAL4; Site_1:
ecori; Site_2: XhoI; ds-cDNA molecules were generated as
selected mRNA by priming with an XhoI oligo-dT primer. The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA polymerase II catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with XhoI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and XhoI sites of the HybridZAP lambda vector
(Stratagene) and excised as pAD-GAL4 phagemids."

BASE COUNT 130 a 168 c 159 g 88 t 1 others
ORIGIN
Query Match 37.7%; Score 255; DB 9; Length 546;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 237 ACCACCACCACCATGTCACACGGCGGCCAGGCTTCGTGTCGCGAGACACGGG 296
DB 217 ACCACCACCACCATGTCACACGGCGGCCAGGCTTCGTGTCGCGAGACACGGG 276
QY 297 TCAGAGGAGCATCAACCTCCACCGCGGAGGTCCACGAGCGAGGAGACTTCCTCG 356
DB 277 TCAGAGGAGCATCAACCTCCACCGCGGAGGTCCACGAGCGAGGAGACTTCCTCG 336
QY 357 CCAGGGCTACTGAGCGCGCGCGCGGATCCACGCCGCTGCTGCTGCTGCTGCTG 416
DB 337 CCAGGGCTACTGAGCGCGCGCGGATCCACGCCGCTGCTGCTGCTGCTGCTGCTG 396
QY 417 CTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
DB 397 CTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 456
QY 477 GCAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
DB 457 GCAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
QY 537 TCCTAA 542
DB 517 TCCTAA 522

RESULT 14
BM501439
LOCUS 417 bp mRNA linear EST 14-FEB-2002
DEFINITION Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
ACCESSION BM501439
VERSION BM501439.1 GI:18661517
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 417)
AUTHORS Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and
Jung,R.
TITLE Maize opaque endosperm mutations create extensive changes in
patterns of gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.

7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.
Location/Qualifiers
1. .417
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI"
BASE COUNT 81 a 150 c 133 g 53 t
ORIGIN

Query Match 35.1%; Score 237; DB 13; Length 417;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 337; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 83 GCACCAATGGCTTACTACAGAGGTGGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
DB 79 GCACCAATGGCTTACTACAGAGGTGGACTGCTCGGAGGAGGTGAGTGGTGGCC 138
QY 143 CCGCGCGCTTCGCGCGCCACGCGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 202
DB 139 CCGCGCGCTTCGCGCGCCACGCGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTC 198
QY 203 GAGGAGGTGACACGGTATCACGCGCGCGGCGCCACCAACCAACCAACCAACCAACCA 262
DB 199 GAGGAGGTGACACGGTATCACGCGCGCGGCGCCACCAACCAACCAACCAACCAACCA 258
QY 263 GCGCGCCACGCTTCGTGCTGCGGAGACGAGGTTCGAGGAGGACATCAACACCTGCACC 322
DB 259 GCGCGCCACGCTTCGTGCTGCGGAGACGAGGTTCGAGGAGGACATCAACACCTGCACC 318
QY 323 GCGGAGGTTCACGAGCGCGGAGGAGCTTCCTCGCCAGGGTAACTAGAGCGCGCGCGGG 382
DB 319 GCGGAGGTTCACGAGCGCGGAGGAGCTTCCTCGCCAGGGTAACTAGAGCGCGCGGG 378
QY 383 CCGGATCATCAGCGCGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
DB 379 CCGGATCATCAGCGCGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417

RESULT 15

AW787315 436 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X2 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787315
VERSION AW787315.1 GI:7844112
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 436)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. .436
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"

```

/clone.lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)".
/tissue.type="tassel, kernal, silk, husk, root, leaf"
/dev.stage="fully-grown"
/lab.host="DHI0B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT      82 a 136 c 136 g 82 t
ORIGIN

```

```

Query Match      34.3%; Score 232; DB 10; Length 436;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 432; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 108 TGGACTACTGTCGGAGGAGTGAGTGGCGCCCGCGGCTTCGGCCGACGCGG 167
Db 1 TGGACTACTGTCGGAGGAGTGAGTGGCGCCCGCGGCTTCGGCCGACGCGG 60

Qy 168 GCGCGCTCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGCG 227
Db 61 GCGCGCTCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGCG 120

Qy 228 CCGCGCCCAACCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGCG 287
Db 121 CCGCGCCCAACCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGCG 180

Qy 288 AGACCCAGGTCGAGGAGGAGTCAACACCTGCACCGCGGAGTTCACGAGCGCAGGAGAG 347
Db 181 AGACCCAGGTCGAGGAGGAGTCAACACCTGCACCGCGGAGTTCACGAGCGCAGGAGAG 240

Qy 348 GCTTCTCGCCAGGCTTAAGTGGTTCGAGGAGTTCACCGCGGAGTTCACGAGCGCAGGAGAG 407
Db 241 GCTTCTCGCCAGGCTTAAGTGGTTCGAGGAGTTCACCGCGGAGTTCACGAGCGCAGGAGAG 300

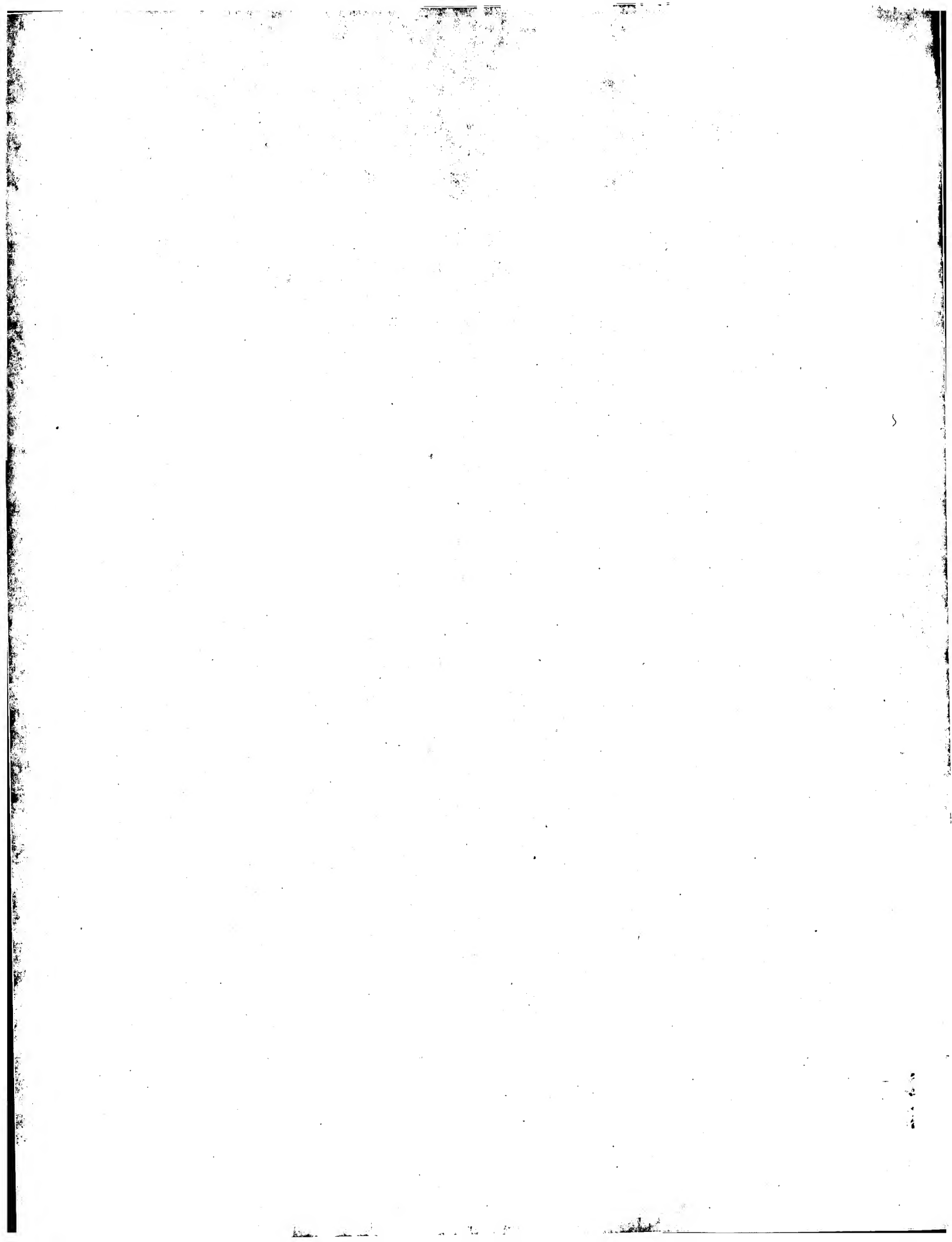
Qy 408 GCCTGGCTGCTTATGTATGTCTGTGTTGACTGGTTCGAGGAGTTCACCGGCTACTTGGCT 467
Db 301 GCCTGGCTGCTTATGTATGTCTGTGTTGACTGGTTCGAGGAGTTCACCGGCTACTTGGCT 360

Qy 468 ATCGTACGTGCACGCGCTCCTGTACGATTAAGAGATTAAGCTCGTACCTGAAT 527
Db 361 ATCGTACGTGCACGCGCTCCTGTACGATTAAGAGATTAAGCTCGTACCTGAAT 420

Qy 528 AAAAATCTCTCGTAAAT 543
Db 421 AAAAATCTCTCGTAAAT 436

```

Search completed: June 16, 2003, 08:38:41
Job time : 1384.39 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:02:07 ; Search time 1293.5 Seconds
(without alignments)
1164.422 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQEVYDCSEVRVAPA.....INTCTGVHERRESFLARAN 93

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+_p2n.model -DEV-xml
-Q/cgn2_1/USPTO_spool/US10090035/runat_06062003_105504_10970/app_query.fasta_1.526
-DB-EST -QFMT-fastap -SUFFIX-rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US10090035 -CGN_1_2463 -runat_06062003_105504_10970 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	417	13	BM501439
2	505	99.0	465	10	BE129897
3	505	99.0	500	10	AW787732
4	505	99.0	523	10	AW331212
5	505	99.0	524	10	AW288876
6	499.5	97.9	466	9	AI964534
C	499.5	97.9	481	9	AI964534
C	499.5	97.9	481	9	AI964534
8	499.5	97.9	539	9	AI854458
9	499.5	97.9	546	9	AA979839
10	499.5	97.9	550	12	BG840383
11	499.5	97.9	553	11	AY104409
12	499.5	97.9	648	14	BQ619167
13	499.5	97.9	648	14	BQ619135
14	499.5	97.9	648	14	BQ619318
15	499.5	97.9	648	14	BQ619337
16	499.5	97.9	648	14	BQ619380
17	499.5	97.9	648	14	BQ619390
18	498	97.6	326	10	BE025303
19	496	97.3	443	10	BE129644
20	496	97.3	552	10	BE025302
21	494	96.9	476	10	BE519299
22	492	96.5	496	10	AW787314
23	482	94.3	433	10	BE225008
24	468	91.8	436	10	AW787315
25	423.5	83.0	399	13	BM318672
26	423.5	83.0	466	10	BE597738
27	423.5	83.0	485	13	BE599123
28	423.5	83.0	513	14	BQ280709
29	423.5	83.0	516	10	BE364814
30	423.5	83.0	519	10	AW680016
31	423.5	83.0	522	10	AW679969
32	423.5	83.0	537	10	AW745400
33	423.5	83.0	541	10	AW745436
34	417.5	81.9	348	10	BE593507
35	410.5	80.5	509	14	BQ280894
36	395	77.5	401	10	AW283056
37	368.5	72.3	274	10	AW288875
38	367.5	72.3	311	12	BF729420
39	308	60.4	453	10	AW923922
40	308	60.4	554	10	AW679915
41	308	60.4	560	10	AW924079
42	308	60.4	566	10	AW677917
43	308	60.4	566	10	AW746383
44	308	60.4	572	10	AW679713
45	308	60.4	572	10	AW925014

ALIGNMENTS

RESULT 1
BM501439
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BM501439
PAC000000000593 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
EST. 14-FEB-2002

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 417)
Hunger, B.G., Beatty, M., Singletary, G., Hamaker, B., Larkins, B.A. and
Jung, R.

TITLE Maize opaque endosperm mutations create extensive changes in
patterns of gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.

FEATURES Location/Qualifiers
Source 1. .417
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI"
BASE COUNT 81 a 150 c 133 g 53 t
ORIGIN

Alignment Scores: 5.6e-47 Length: 417
Pred. No.: 510.00 Matches: 93
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 13 Gaps: 0
DB: 13

US-10-090-035-4 (1-93) x BM501439 (1-417)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProAla 20
Db 85 ATGGCTTACTACGAGGAGTGGACTACTGCTCGAGGAGTGAGTGGTGGCCCGGCC 144
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 145 GGCTTCGGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 204
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisGly 60
Db 205 GTCCACACGGTCTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 264
QY 61 HisGlyPheValValArgGluThrArgValGluAspPheAsnThrCysThrGlyGlu 80
Db 265 CACGGCTTCTGTGTGCGGAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 324
QY 81 ValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 325 GTCCACGAGCGCAGGAGAGGCTTCTCGCCAGGCGCTAAC 363

RESULT 2
BE129897 465 bp mRNA linear EST 21-JUN-2000
LOCUS 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897.1 GI:8577260
VERSION BE129897.1
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 465)

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Contact: Walbot V
Unpublished (1999)
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

FEATURES
source

Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1. .465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="PH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 88 a 151 c 149 g 77 t
ORIGIN

Alignment Scores: 2.26e-46 Length: 465
Pred. No.: 505.00 Matches: 92
Score: 98.92% Conservative: 0
Percent Similarity: 98.92% Mismatches: 1
Best Local Similarity: 98.92% Indels: 0
Query Match: 10 Gaps: 0
DB: 10

US-10-090-035-4 (1-93) x BE129897 (1-465)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProAla 20
Db 40 ATGGCTTACTACGAGGAGTGGACTACTGCTCGAGGAGTGAGTGGTGGCCCGGCC 99
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 100 GGCTTCGGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 159
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisGly 60
Db 160 GTCCACACGGTCTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 219
QY 61 HisGlyPheValValArgGluThrArgValGluAspPheAsnThrCysThrGlyGlu 80
Db 220 CACGGCTTCTGTGTGCGGAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 279
QY 81 ValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 280 TTCCACGAGCGCAGGAGAGGCTTCTCGCCAGGCGCTAAC 318

RESULT 3
AW787732 500 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW787732.1 GI:7844510
VERSION AW787732
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 500)

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Contact: Walbot V
Unpublished (1999)
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 945002 row: E column: 06.

FEATURES

source
 1..500
 Location/Qualifiers
 /organism="Zea mays"
 /cultivar="W23"
 /clone_lib="945" - Mixed adult tissues from Walbot lab,
 same as 707 (SK)
 /db_xref="taxon:4577"
 /tissue_type="tassel, kernel, silk, husk, root, leaf"
 /dev_stage="fully-grown"
 /lab_host="DH108"
 /note="Organ: tassel, kernel, silk, husk, root, leaf;
 Vector: pGAD10; Site1: EcoRI; cDNA library from fully
 differentiated maize tissues from an active Mutator plant.
 Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
 root, leaf). Unidirectionally cloned. New library number
 given to library 707 for additional sequencing."
 97 a 159 c 156 g 88 t

BASE COUNT

ORIGIN
 Alignment Scores:
 Pred. No.: 2,45e-46 Length: 500
 Score: 505.00 Matches: 92
 Percent Similarity: 98.92% Conservative: 0
 Best Local Similarity: 98.92% Mismatches: 1
 Query Match: 99.02% Indels: 0
 DB: 10 Gaps: 0

US-10-090-035-4 (1-93) x AW787732 (1-500)

QY 1 MetaLysTyrrClnGluValAspTyrCysSerGluValArgSerValAlaProAla 20
 Db ATGGCTTACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTGGCCCGGCC 94
 QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 Db GGCTTCGGCCGCGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAG 154
 QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGlyGly 60
 Db GTGCACACGGTCTACGCGCGCGCGCCCAACCAACCAACCAACCAACCAACCAACCA 214
 QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
 Db CACGGCTTGGTGGTGGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGGAG 274
 QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 Db TTCCACGAGCGCGAGAGGCTTCTCCCGAGGCTTAAC 313

RESULT 4

AW331212
 LOCUS 707049E04.x1 707 - Mixed adult tissues from Walbot lab (SK) Zea
 DEFINITION mays cDNA, mRNA sequence.
 ACCESSION AW331212
 VERSION AW331212.1 GI:6827569
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 523)

Walbot, V.
 Zea mays ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 JOURNAL
 COMMENT Contact: Walbot V
 Department of Biological Sciences

Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 707049 row: E column: 04.

FEATURES

source
 1..523
 Location/Qualifiers
 /organism="Zea mays"
 /cultivar="W23"
 /db_xref="taxon:4577"
 /clone_lib="707" - Mixed adult tissues from Walbot lab (SK)
 /tissue_type="tassel, kernel, silk, husk, root, leaf"
 /dev_stage="adult"
 /lab_host="DH108"
 /note="Organ: tassel, kernel, silk, husk, root, leaf;
 Vector: pGAD10; Site1: EcoRI; cDNA library from fully
 differentiated maize tissues from an active Mutator
 plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
 husk, root, leaf). Unidirectionally cloned."
 125 a 158 c 149 g 91 t

BASE COUNT

ORIGIN
 Alignment Scores:
 Pred. No.: 2,58e-46 Length: 523
 Score: 505.00 Matches: 92
 Percent Similarity: 98.92% Conservative: 0
 Best Local Similarity: 98.92% Mismatches: 1
 Query Match: 99.02% Indels: 0
 DB: 10 Gaps: 0

US-10-090-035-4 (1-93) x AW331212 (1-523)

QY 1 MetaLysTyrrClnGluValAspTyrCysSerGluValArgSerValAlaProAla 20
 Db ATGGCTTACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTGGCCCGGCC 94
 QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 Db GGCTTCGGCCGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAG 154
 QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGlyGly 60
 Db GTGCACACGGTCTACGCGCGCGCGCCCAACCAACCAACCAACCAACCAACCAACCA 214
 QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
 Db CACGGCTTGGTGGTGGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGGAG 274
 QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 Db TTCCACGAGCGCGAGAGGCTTCTCCCGAGGCTTAAC 313

RESULT 5

AW288876
 LOCUS 707009E07.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea
 DEFINITION mays cDNA, mRNA sequence.
 ACCESSION AW288876
 VERSION AW288876.1 GI:6695663
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 524)

Walbot, V.
 Zea mays ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 JOURNAL
 COMMENT Contact: Walbot V

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES

Location/Qualifiers
1. 524
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707" - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="PH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site: 1; EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
107 a 170 c 157 g 89 t 1 others

BASE COUNT

ORIGIN
Alignment Scores:
Pred. No.: 2.58e-46 Length: 524
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-4 (1-93) x AW288876 (1-524)

QY 1 MetAlaTyTyTGlnGluValAlaSPtyrCysSerGluGluValArgSerValAlaProAla 20
DB 73 ATGGCTTACTACCGAGGTTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGCC 132
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 133 GGCCTTCGGCGCCACGCGCGCGGTCCAGCAGCACGTCGTCGAAGAGAGTTCGAGGAG 192
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisGly 60
DB 193 GTCGACAGGTCCTACGCGCGCGCGCCACACACACCATGTCACACCGCGCGC 252
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 253 CACGGCTTCGTGTCGCGAGACACGAGGTGCAAGAGGACATCAACACCTGCACCGCGGAG 312
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 313 TTCCACGAGCGGAGGAGAGGCTTCTCGCCAGGCGCTAAC 351

RESULT 6
AI964534/c 466 bp mRNA linear EST 20-AUG-1999
LOCUS 496013D01.x1 496 - stressed shoot cDNA library from Wang/Bohnert
DEFINITION lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964534
VERSION AI964534.1 GI:5757247
KEYWORDS EST.

SOURCE

ORGANISM Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 466)

REFERENCE

AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
Unpublished (1999)

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496013 row: D column: 01.

FEATURES

Location/Qualifiers
1. 466
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496" - stressed shoot cDNA library from
Wang/Bohnert lab
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+) XR;
Wang/Bohnert"
76 a 151 c 154 g 85 t

BASE COUNT

ORIGIN
Alignment Scores:
Pred. No.: 9.2e-46 Length: 466
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AI964534 (1-466)

QY 1 MetAlaTyTyTGlnGluValAlaSPtyrCysSerGluGluValArgSerValAlaProAla 20
DB 416 ATGGCTTACTACCGAGGTTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGCC 357
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 356 GGCCTTCGGCGCCACGCGCGCGGTCCAGCAGCACGTCGTCGAAGAGAGTTCGAGGAG 297
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisHisGly 59
DB 296 GTCGACAGGTCCTACGCGCGCGCCACACACCATGTCACACCGCGCGC 237
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
DB 236 GGCACCGGCTTCGTGTCGCGAGACACGAGGTGCGAGGAGGACATCAACACCTGCACCGC 177
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
DB 176 GAGGTCCAGGAGCGAGGAGAGTTCCTCGCCAGGCGTAAAC 135

RESULT 7

AI964458/c 481 bp mRNA linear EST 20-AUG-1999
LOCUS 496010H04.x1 496 - stressed shoot cDNA library from Wang/Bohnert
DEFINITION lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964458
VERSION AI964458.1 GI:5757171
KEYWORDS EST.

SOURCE

ORGANISM Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 481)

REFERENCE

AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
Unpublished (1999)
CONTACT: Walbot V
Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Email: walbot@stanford.edu
Plate: 496010 row: K column: 04.

FEATURES	SOURCE
-----------------	---------------

```

location/Qualifiers
1..481
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from
wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+) XbaI"

```

wang/Bonhert"		
BASE COUNT	82 a	148 c
ORIGIN	154 g	97 t

Alignment Scores:		
Pred. No.:	9.53e-46	481
Score:	499.50	93
Percent Similarity:	98.94%	0
Best Local Similarity:	98.94%	Conservative: 0
Query Match:	97.94%	Mismatches: 0
DB:	9	Indels: 1
		Gaps: 1

US-10-090-035-4 (1-93) x AI964458 (1-481)

QY	1	MetAlaTyr	Tyr	Gln	Glu	Val	Asp	Tyr	Cys	Ser	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20
Db	451	ATGCGTTACT	ACC	AGG	AGT	GGACTACT	CTCG	GAG	GAG	TGAGGT	CGTGGCGCCCGCGGC	392							
QY	21	GlyPhe	Gly	Arg	His	Gly	Gly	Val	Gln	Gln	His	Val	Val	Lys	Glu	Lys	Phe	Glu	40
Db	391	GGCTTCGGCC	CCAC	GAG	GCGGGCTCC	ACGACGTCGTC	AAG	GAG	AAG	TTCG	AGAG	332							
QY	41	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	---	His	His	His	His	Gly	His	59
Db	331	GTCCACAC	CGGTCT	CAC	GCGCGCGGCC	CAAC	CAC	CAC	CAC	CAC	CAC	CGT	TGTC	CAC	CAC	CGC	272		
QY	60	Gly	His	Gly	Phe	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Asp	Phe	Asn	Thr	Cys	79
Db	271	GGCCAC	CGCTCT	TGTTG	TCGCG	GAGAC	CGGTTC	GAG	GAG	ACATCA	ACACCT	GCAC	CGC	212					
QY	80	Glu	Val	His	Glu	Arg	Glu	Ser	Phe	Leu	Ala	Arg	Ala	Asn	93				
Db	211	GAGTTC	CAC	GAG	CGG	AGAG	CGTTCCT	CGC	CAG	GGCTAAC	170								

RESULT 8
AT855425/C

AI855425 539 bp mRNA linear EST 15-JUL-1999
603016F02.x1 603 - stressed root cDNA library from Wang/Bohnert lab
Zea mays cDNA, mRNA sequence.

ACCESSION
VERSION

KEYWORDS EST.
SOURCE Zea mays.

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 539)

AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL
University
Unpublished (1999)

COMMENT
Contact: Walbot V
Department of Biological Sciences

Department of Biological Sciences
Stanford University
855 California Ave.
Stanford, CA 94305-5080

655 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.

	FEATURES	SOURCE
1.	100% cotton	100%
2.	100% polyester	100%
3.	100% rayon	100%
4.	100% silk	100%
5.	100% wool	100%
6.	100% linen	100%
7.	100% leather	100%
8.	100% rubber	100%
9.	100% metal	100%
10.	100% plastic	100%
11.	100% glass	100%
12.	100% paper	100%
13.	100% wood	100%
14.	100% stone	100%
15.	100% concrete	100%
16.	100% brick	100%
17.	100% tile	100%
18.	100% paint	100%
19.	100% ink	100%
20.	100% glue	100%
21.	100% oil	100%
22.	100% water	100%
23.	100% air	100%
24.	100% fire	100%
25.	100% electricity	100%
26.	100% magnetism	100%
27.	100% gravity	100%
28.	100% light	100%
29.	100% sound	100%
30.	100% heat	100%
31.	100% cold	100%
32.	100% dryness	100%
33.	100% wetness	100%
34.	100% softness	100%
35.	100% hardness	100%
36.	100% smoothness	100%
37.	100% roughness	100%
38.	100% clean	100%
39.	100% dirty	100%
40.	100% new	100%
41.	100% old	100%
42.	100% young	100%
43.	100% adult	100%
44.	100% child	100%
45.	100% man	100%
46.	100% woman	100%
47.	100% boy	100%
48.	100% girl	100%
49.	100% father	100%
50.	100% mother	100%
51.	100% brother	100%
52.	100% sister	100%
53.	100% uncle	100%
54.	100% aunt	100%
55.	100% grandfather	100%
56.	100% grandmother	100%
57.	100% great-grandfather	100%
58.	100% great-grandmother	100%
59.	100% great-great-grandfather	100%
60.	100% great-great-grandmother	100%
61.	100% great-great-great-grandfather	100%
62.	100% great-great-great-grandmother	100%
63.	100% great-great-great-great-grandfather	100%
64.	100% great-great-great-great-grandmother	100%
65.	100% great-great-great-great-great-grandfather	100%
66.	100% great-great-great-great-great-grandmother	100%
67.	100% great-great-great-great-great-great-grandfather	100%
68.	100% great-great-great-great-great-great-grandmother	100%
69.	100% great-great-great-great-great-great-great-grandfather	100%
70.	100% great-great-great-great-great-great-great-grandmother	100%
71.	100% great-great-great-great-great-great-great-great-grandfather	100%
72.	100% great-great-great-great-great-great-great-great-grandmother	100%
73.	100% great-great-great-great-great-great-great-great-great-grandfather	100%
74.	100% great-great-great-great-great-great-great-great-great-grandmother	100%
75.	100% great-great-great-great-great-great-great-great-great-great-grandfather	100%
76.	100% great-great-great-great-great-great-great-great-great-great-grandmother	100%
77.	100% great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
78.	100% great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
79.	100% great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
80.	100% great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
81.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
82.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
83.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
84.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
85.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
86.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
87.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
88.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
89.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
90.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
91.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
92.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
93.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandfather	100%
94.	100% great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-great-grandmother	100%
95.	100% great-grandfather	100%
96.	100% great-grandmother	100%
97.	100% great-grandfather	100%
98.	100% great-grandmother	100%
99.	100% great-grandfather	100%
100.	100% great-great-great-g	

```

/organism="Zea mays"
/contig="B73"
/db_xref="taxon:4577"
/clone_lib="603" stressed root cDNA library from
wang/bonner lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/name="E. coli XL Gold"
/notes="Organ: root; Vector:
pBluescriptII SK(+)"

```

BASE COUNT	91 a	157 c	172 g	119 t	seedling stressed root cDNA library from Wang/Bohnert lab
ORIGIN					

Alignment Scores:

Pred. No.:	1.08e-45	Length:	539
Score:	499.50	Matches:	93
Percent Similarity:	98.94%	Conservative:	0
Best Local Similarity:	98.94%	Mismatches:	0
Query Match:	97.94%	Indels:	1
DB:	9	Gaps:	1

US-10-090-035-4 (1-93) x AI855425 (1-539)

QY	1	Meta	Ty	Tyr	Gln	Glu	Val	Asp	Tyr	Cys	Ser	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20												
DB	477	AT	GG	CTT	TACT	AC	CAG	GAG	TGG	ACT	TG	TC	CG	SAG	GAG	TG	AGG	TG	CGT	GGC	CGCC	418										
QY	21	Gly	Phe	Gly	Arg	His	Gly	Gly	Val	Gln	His	Val	Val	Lys	Glu	Lys	Phe	Glu	Glu	40												
DB	417	GG	CT	TC	GG	CC	CC	CAC	GAG	GG	CG	CT	CC	AG	CAC	GC	TC	TA	AG	GAG	AA	GT	TC	CG	AG	AG	358					
QY	41	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	---	His	His	His	His	His	Gly	His	His	Gly	59										
DB	357	GT	CAC	AC	CG	GT	CT	CAC	GC	GC	GG	CG	CC	AA	CAC	CAC	CA	CC	AC	CA	CC	AT	GG	T	CAC	CA	CG	298				
QY	60	Gly	His	Gly	Phe	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Asp	Phe	Asn	Thr	Cys	Thr	Gly	79											
DB	297	GG	CC	CAC	GG	CT	CT	CG	GT	GT	CC	CG	AG	ACC	AG	CC	GG	TC	CG	AG	GAG	CA	TC	AA	CA	CC	TC	GC	AC	CG	CG	238
QY	80	Glu	Val	His	Glu	Arg	Arg	Glu	Ser	Phe	Leu	Ala	Arg	Ala	Asn	93																
DB	237	GAG	GT	CAC	CAG	CG	CG	AGG	AG	AG	CT	TC	CT	CG	CC	AGG	CG	CT	TAAC	196												

RESULT 9

AA979839 LOCUS 546 bp mRNA linear EST 26-MAY-1998
MEST2-B7.1.TW1412.Seg ISUM2 Zea mays cDNA clone MEST2-B7.5, mRNA
sequence.
AA979839 ACCESSION
AA979839.1 GI:3157217
EST.
KEYWORDS
Zea mays.
SOURCE

ORGANISM Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE	1 (bases 1 to 546)
AUTHORS	Wen, T. J., Ashlock, D. A. and Schnable, P. S.
TITLE	Expressed Sequence Tags from B73 Maize Seedlings
JOURNAL	Unpublished (1997)
COMMENT	Contact: Schnable, ps

Contact: Schnable, PS
Schnable Laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975
Fax: (515)-294-2299
Email: schnable@iastate.edu

Mon Jun 16 14:55:59 2003

AUTHORS TITLE JOURNAL COMMENT

Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206705.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

PCR Primers

FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T7-1 (AA TAC GAC CTA TAG)

FEATURES

source

Location/Qualifiers

1..550

/organism="Zea mays"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="MEST12-H11"

/clone_lib="ISUM4-TN"

/tissue_type="Seedling and silk"

/lab_host="DH10B"

/note="vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT, selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGAAGATTTCGGCGCGAGGATTTTTTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT7T3PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

Research 6: 791-806, 1996."

BASE COUNT 124 a 174 c 159 g 92 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1..11e-45 Length: 550

Score: 499.50 Matches: 93

Percent Similarity: 98.94% Conservative: 0

Best Local Similarity: 98.94% Mismatches: 0

Query Match: 97.94% Indels: 1

DB: 12 Gaps: 1

US-10-090-035-4 (1-93) x BG840383 (1-550)

QY 1 MetaLalTyrrGluValAlaSerCysSerGluGluValAlaProAla 20

DB 75 ATGGCTTACTACGAGAGGTGGACACCTCGCGAGGAGGTGAGTCCGCGCGGCC 134

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40

DB 135 GGCTTCGGCGCGCCACGAGCGCGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 194

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59

DB 195 GTCGACACGGTCTCAGCGCGCGCGCCACACACACACACACACACATGGTCACCGCGC 254

QY 60 GlyHisGlyPheValValAlaArgGluThrArgValGluAlaAspIleAsnThrCysThrGly 79

DB 255 GGCCACGGCTTCGTGGTGGCGGAGACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 314

QY 80 GluValHisGluAlaArgGluSerPheLeuAlaArgAlaAsn 93

DB 315 GAGGTCCACGAGCGAGGAGAGCTTCCTCGCCAGGCGCTAAC 356

RESULT 11

AY104409

PCR Primers

FORWARD: T71412 (5'-GAAGATACCCCAACCAACC-3')
BACKWARD: T71412 (5'-TAATAGCACTCATATAGGC-3')
Plate: MEST2 row: B column: 7
Seq primer: tw1412 (5'-GAAGATACCCCAACCAACC-3')

FEATURES

source

Location/Qualifiers

1..546

/organism="Zea mays"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="MEST2-B7"

/clone_lib="ISUM2"

/tissue_type="above ground tissues"

/dev_stage="Two-leaf-stage green seedling"

/lab_host="XLI-MFR Blue"

/note="Organ: green seedlings; Vector: pAD-GAL4; Site 1: EcoRI; Site 2: XhoI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with an XhoI oligo-dT primer. The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with XhoI and size-selected. The resulting molecules were directionally cloned into the EcoRI and XhoI sites of the HybriZAP lambda vector (Stratagene) and excised as pAD-GAL4 phagemids."

BASE COUNT 130 a 168 c 159 g 88 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1..1e-45 Length: 546

Score: 499.50 Matches: 93

Percent Similarity: 98.94% Conservative: 0

Best Local Similarity: 98.94% Mismatches: 0

Query Match: 97.94% Indels: 1

DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AA979839 (1-546)

QY 1 MetaLalTyrrGluValAlaSerCysSerGluGluValAlaProAla 20

DB 66 ATGGCTTACTACGAGAGGTGGACACCTCGCGAGGAGGTGAGTCCGCGCGGCC 125

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40

DB 126 GGCTTCGGCGCGCCACGAGCGCGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 185

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59

DB 186 GTCGACACGGTCTCAGCGCGCGCGCCACACACACACACACACATGGTCACCGCGC 245

QY 60 GlyHisGlyPheValValAlaArgGluThrArgValGluGluAlaAspIleAsnThrCysThrGly 79

DB 246 GGCCACGGCTTCGTGGTGGCGGAGACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 305

QY 80 GluValHisGluAlaArgGluSerPheLeuAlaArgAlaAsn 93

DB 306 GAGGTCCACGAGCGAGGAGAGCTTCCTCGCCAGGCGCTAAC 347

RESULT 10

BG840383

LOCUS

DEFINITION

MEST12-H11.77-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA

sequence.

ACCESSION

BG840383

VERSION

BG840383.2

GI:14242676

KEYWORDS

EST.

SOURCE

Zea mays.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 550)

LOCUS AY104409 553 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PC0124784 mRNA sequence.
 ACCESSION AY104409
 VERSION AY104409.1 GI:21207487
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays

REFERENCE 1 (bases 1 to 553)
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 553)
 AUTHORS Coe,E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA

FEATURES
 source
 1..553
 Location/Qualifiers
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone_lib="PC0124784"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

BASE COUNT 128 a 173 c 162 g 90 t

Alignment Scores:
 Pred. No.: 1.11e-45 Length: 553
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 11 Gaps: 1

US-10-090-035-4 (1-93) x AY104409 (1-553)

Qy 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 76 ATGGCTTACTACCGAGGAGTGGACTACTGCTCGAGAGGAGTGGAGTGGCGCCGGCC 135

Qy 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
 |||||
 Db 136 GGCTTCGGCCGCCACGAGGCGGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAG 195

Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisHisGly 59
 |||||
 Db 196 GTCGACACGGTCTCAGCGCCGCCGCCACCCACCACCATGTCACACGCGC 255

Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
 |||||
 Db 256 GGCCACGGTCTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 315

Qy 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 316 GAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 357

RESULT 12
 BQ619167 648 bp mRNA linear EST 27-JUN-2002
 LOCUS RNOSEQ4E05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
 DEFINITION mays cDNA clone RNOSEQ4E05_SK.ab1 similar to No homology, mRNA

ACCESSION BQ619167
 VERSION RNOSEQ4E05_SK.ab1
 KEYWORDS Zea mays.
 SOURCE Zea mays

sequence.
 BQ619167
 VERSION BQ619167.1 GI:21621161
 KEYWORDS Zea mays.
 SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 648)
 AUTHORS Wang,H. and Bohnert,H.J.
 TITLE Genomics of plant stress tolerance
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mark Fredricksen
 Department of Plant Biology
 University of Illinois
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 2172653473
 Email: bohnertlablife.uiuc.edu.

FEATURES
 source
 1..648
 Location/Qualifiers
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone_lib="RNOSEQ4E05_SK.ab1"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150
 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

Alignment Scores:
 Pred. No.: 1.33e-45 Length: 648
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x BQ619167 (1-648)

Qy 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 82 ATGGCTTACTACCGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGCGCCGGCC 141

Qy 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
 |||||
 Db 142 GGCTTCGGCCGCCACGAGGCGGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAG 201

Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisHisGly 59
 |||||
 Db 202 GTCGACACGGTCTCAGCGCCGCCGCCACCCACCACCATGTCACACGCGC 261

Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
 |||||
 Db 262 GGCCACGGTCTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321

Qy 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 322 GAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 363

RESULT 13
 BQ619315 648 bp mRNA linear EST 27-JUN-2002
 LOCUS RNOSEQ6C06_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
 DEFINITION mays cDNA clone RNOSEQ6C06_SK.ab1 similar to No homology, mRNA

ACCESSION BQ619315
 VERSION BQ619315.1 GI:21621309
 KEYWORDS Zea mays.
 SOURCE Zea mays

100

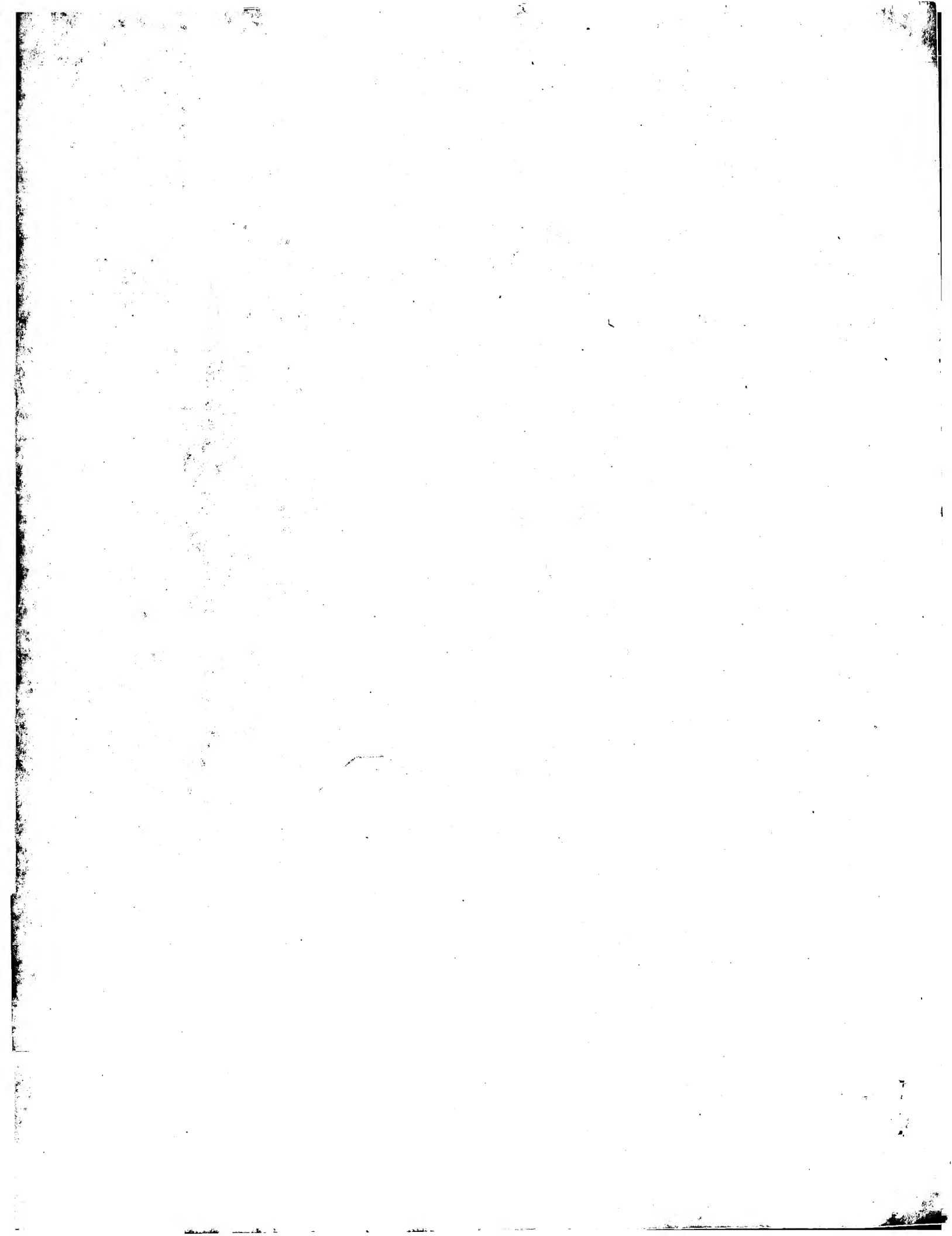
FEATURES Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
source 1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6E12_SK.ab1"
/tissue_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1,33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x BQ619337 (1-648)

OY 1 MetalaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValalaProAla 20
|||||
Db 82 ATGGCTTACTACGAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGGCCCCGGCC 141
OY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 142 GGCTTCGGCCGCCACGGAGCGGGGTCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201
OY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisHisGly 59
|||||
Db 202 GTCGACACGGTCTCAACGGCCCGCGCCACCCACCACCACCATGGTCAACACGGC 261
OY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||
Db 262 GGCACCGGCTTCGTGGTGGCGAGACCGAGGTCGAGGAGGAGACATCAACACCTGCACCGGC 321
OY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 322 GAGTCCACGAGCGCAGGAGAGGCTTCCTCGCCAGGCGCTAAC 363

Search completed: June 16, 2003, 12:16:36
Job time : 1296.5 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 01:23:01 ; Search time 1222.39 seconds
(without alignments)
7604.943 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.2	90.8	553	11	AY104409
2	518	90.2	552	10	BE025302
3	509.6	88.8	546	9	AA979839
4	507.2	88.4	524	10	AW288876
5	500.8	87.2	523	10	AW331212
6	500.6	87.2	648	14	BE0619167

7	500.6	87.2	648	14	BQ619315
8	500.6	87.2	648	14	BQ619318
9	500.6	87.2	648	14	BQ619337
10	500.6	87.2	648	14	BQ619383
11	500.6	87.2	648	14	BQ619390
12	494.6	86.2	550	12	BG840383
13	494.4	86.1	539	9	AI855425
14	485.8	84.6	500	10	AW787732
15	465.6	81.1	496	10	AW787314
16	461.8	80.5	465	10	BE129897
17	457.2	79.7	476	10	BE129897
18	455.4	79.3	481	9	AI851929
19	440.4	76.7	466	9	AI964534
20	431.2	75.1	436	10	AW787315
21	426.2	74.3	443	10	BE129644
22	417	72.6	417	13	BM501439
23	411.4	71.7	433	10	BE225008
24	362.8	63.2	401	10	AW289056
25	326.6	56.9	541	10	AW745436
26	324.8	56.6	537	10	AW745400
27	322.8	56.2	326	10	BE025303
28	308.2	53.7	516	10	BE364814
29	308	53.7	519	10	AW680016
30	307.2	53.5	485	10	BE599423
31	307.2	53.5	513	14	BQ280709
32	304	53.0	522	10	AW679969
33	303.6	52.9	509	14	BQ280894
34	291.4	50.8	466	10	BE597738
35	274	47.7	362	12	BG840882
36	272.2	47.4	311	12	BF729420
37	251.4	43.8	274	10	AW288875
38	242	42.2	367	10	AW747603
39	240.6	41.9	399	13	BM318672
40	220.4	38.4	348	10	BE593507
41	219.6	38.3	226	9	AI372183
42	211	36.8	288	9	AI649804
43	208.8	36.4	338	10	AW924579
44	186	32.4	240	14	W21677
45	185.4	32.3	512	14	BQ283209

ALIGNMENTS

RESULT 1
AY104409
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
Clade; Panicoideae; Andropogoneae; Zea.
Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 553)
Coe,E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
Location/Qualifiers
1..553
/organism="Zea mays"
/db_xref="MaizeDB:637271"
/db_xref="taxon:4577"
/clone="PC0124784"

linear HTC 25-MAY-2002

Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945028 row: B column: 09.
Location/Qualifiers
1. 552
/organism="Zea mays"
/cultivar="W23"
/db.xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 135 a 165 c 160 g 92 t
ORIGIN

Query Match 90.2%; Score 518; DB 10; Length 552;
Best Local Similarity 99.0%; Pred. No. 5.1e-68;
Matches 521; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

49 CGACGTACACGGGGCTGGCGACAGACACACCAAGCGTCGGCACCACCAATGCTTACTACC 108
1 CGACGTACACGGGGCTGGCGACAGACACCAAGCGTCGGCACCACCAATGCTTACTACC 60
109 AGGAGGTGACGTACTGCTCGGAGGAGGTGAGTGTGCTGGCGCCCGCGCTTCGGCCGCC 168
61 AGGAGGTGACGTACTGCTCGGAGGAGGTGAGTGTGCTGGCGCCCGCGCTTCGGCCGCC 120
169 AGGGCGGGCGGCTCCAGCAGACAGTGTCAAGGAGAGTTCGAGGAGGTTCGACAGGCT 228
121 AGGGCGGGCGGCTCCAGCAGACAGTGTCAAGGAGAGTTCGAGGAGGTTCGACAGGCT 180
229 CAGCGCGCGCGCCAC 288
181 CAGCGCGCGCGCCAC 240
289 TGGCGGAGACACAGGCTCGAAGAGGACATCAACACCTGCACCGCGGAGGTTCACGAGCGCA 348
241 TGGCGGAGACACAGGCTCGAAGAGGACATCAACACCTGCACCGCGGAGGTTCACGAGCGCA 300
349 GGGAGAGTTCCTCGCCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 408
301 GGGAGAGTTCCTCGCCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 360
409 GTGCTTGCCTGCGCTTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 468
361 GTGCTTGCCTGCGCTTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 420
469 TTGGCTATCTAGTGTGACCGACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
421 TTGGCTATCTAGTGTGACCGACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
529 CTGAATAAACTTCTTCTGTAATACTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 574
481 CTGAATAAACTTCTTCTGTAATACTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 526

Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945028 row: B column: 09.
Location/Qualifiers
1. 552
/organism="Zea mays"
/cultivar="W23"
/db.xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 128 a 173 c 162 g 90 t
ORIGIN

Query Match 90.8%; Score 521.2; DB 11; Length 553;
Best Local Similarity 98.0%; Pred. No. 1.7e-68;
Matches 539; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

24 CGCACAGCAATCCACACAGCACTTCGACGTCCACAGCGGCGTGGCGACAGACACACCAA 83
4 CGACGACCAATCCACACAGCACTTCGACGTCCACAGCGGCGTGGCGACAGACACACCAA 63
84 CGCTGCGCACAATGGCTTACTACAGAGGTGGACTACTGTCTGGAGGAGGTGAGTGG 143
64 CGCTGCGCACAATGGCTTACTACAGAGGTGGACTACTGTCTGGAGGAGGTGAGTGG 123
144 GTGGCCCCCGCGCTTCGCCGCCACAGCGCGGCTCCAGCAGACAGTGTCAAGGAG 203
124 GTGGCCCCCGCGCTTCGCCGCCACAGCGCGGCTCCAGCAGACAGTGTCAAGGAG 183
204 AAGTTCGAGGAGTTCGACACAGGCTCTCACGCGCGCGGCCA---ACACACACACACCACT 260
184 AAGTTCGAGGAGTTCGACACAGGCTCTCACGCGCGCGGCCA---ACACACACACACCACT 243
261 GGTACACACGGCGCCACGCTTCGTTGGTGGCGAGACAGGCTCGAAGGAGACATCAAC 320
244 GGTACACACGGCGCCACGCTTCGTTGGTGGCGAGACAGGCTCGAAGGAGACATCAAC 303
371 ACCTGCACCGCGAGGTCACAGAGCGCAGGAGAGTTCCTCCGCGAGGCTAACTGAGCC 380
304 ACCTGCACCGCGAGGTCACAGAGCGCAGGAGAGTTCCTCCGCGAGGCTAACTGAGCC 363
381 GCCCGGGCGCGGATCCACGCGCGTTCGTTGGTGGCGAGACAGGCTCGAAGGAGACATCA 440
364 GCCCGGGCGCGGATCCACGCGCGTTCGTTGGTGGCGAGACAGGCTCGAAGGAGACATCA 423
441 GTTGACTGTTGTCAGGCTCATCGTACTTGGCTATCGTACGTGACGCGACTCAGCTCCT 500
424 GTTGACTGTTGTCAGGCTCATCGTACTTGGCTATCGTACGTGACGCGACTCAGCTCCT 483
501 GTACGAATTACGAATAAGCTGTGACCTGAATAAATCTTCGTAATACTAAAAA 560
484 GTACGAATTACGAATAAGCTGTGACCTGAATAAATCTTCGTAATACTAAAAA 543
561 AAAAAA 570
544 AAAAAA 553

RESULT 2
LOCUS BE025302
DEFINITION 945028B09.Y1 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE025302
VERSION BE025302.1 GI:8318737
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 552)
REFERENCE Walbot,V.
AUTHORS Zea mays ESTs from various cDNA libraries sequenced at Stanford University
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University

AA979839 546 bp mRNA linear EST 26-MAY-1998
LOCUS AA979839
DEFINITION MEST2-B7.TW1412.Seq ISM2 Zea mays cDNA clone MEST2-B7 5', mRNA


```

/clone="RNSEQ4E05_SK.abl"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"

```

BASE COUNT	172 a	193 c	183 g	100 t
ORIGIN				
Query Match		87.2%	Score 500.6;	DB 14;
Best Local Similarity		97.6%	Pred. No. 1.8e-05;	Length 648;
Matches 530; Conservative		0;	Mismatches	0; Indels 4; Gaps 2;

36	QY	CACACAGCAGCTTCGACGTCCACACGGG-CGCTGCGCACAGACACACCAAGCGTGGCACC	94
21	Db	CACGAGACAGCTTCGACGTTCGACGGCGCGTSCACACAGACACACCAAGCGTGGCACC	80
95	QY	AATGCGTTACTACACAGAGGTGAGCTACTGCTCGGAGGAGGTGAGGTGCGGCCCGGC	154
81	Db	AATGCGTTACTACACGAGGTGAGCTACTGCTCGGAGGAGTGAGTCCGTGGCCCGCGC	140
155	QY	CGGCTTCGGCGCCACCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGA	214
141	Db	CGGCTTCGGCGCCACCGAGGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGA	200
215	QY	GGTCGACACGGTCTCAGCGCCCGCGGCCCA---ACCACCAACCACCACTGTCACCAACGG	271
201	Db	GGTCGACACGGTCTCAGCGCCCGCGCGCAACCAACCACCACTGTCACCAACGG	260
272	QY	CGGCCACGGCTTGGTGGCGGAGACCAAGGCTCGAAGGACATCAACACCTCGCACCGG	331
261	Db	CGGCCACGGCTTGGTGGCGGAGACCAAGGCTCGAGGAGCATCAACACCTCGCACCGG	320
332	QY	CGAGGTCACGACGCGCAGGAGAGCTTCCTCGCCAGGCGCTAACTGAGCGCCCGCGCGCC	391
321	Db	CGAGGTCACGACGCGCAGGAGAGCTTCCTCGCCAGGCGCTAACTGAGCGCGCGCGCGCC	380
392	QY	GGCATCCACCGCGGTCGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGT	451
381	Db	GGCATCCACCGCGGTCGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGT	440
452	QY	GTTACAGGTCATCGTACTTGGCTATCGTACGTGCAAGCACTCAGCTCCTGTACGAATTAC	511
441	Db	GTGCAAGGTCATCGTACTTGGCTATCGTACGTGCAAGCACTCAGCTCCTGTACGAATTAC	500
512	QY	GACAATAAGCTCGTGAACCTGAATAAACTTCTTCGTAATACTAAAAAATAAAAAAAA	571
501	Db	GACAATAAGCTCGTGAACCTGAATAAACTTCTTCGTAATACTAAAAAATAAAAAAAA	560
572	QY	AAA	574
561	Db	AAA	563

RESULT 7
 B0619315
 LOCUS
 DEFINITION
 B0619315 648 bp mRNA linear EST 27-JUN-2002
 RN05Q6C06_SK.ab1 Salt stressed zea mays roots cDNA library zea
 mays cDNA clone RN05Q6C06_SK.ab1 similar to No homology, mRNA
 sequence.
 B0619315
 B0619315.1 GI:21621309
 EST.
 zea mays.
 zea mays
 zea mays
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 648)
 Wang, H. and Bohnert, H.J.
 Genomics of plant stress tolerance
 Unpublished (2002)
 Contact: Mark Fredricksen

Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217/2655473
Email: bohnertlab@life.uiuc.edu.

```

FEATURES
  source
    Location/Qualifiers
      1..548
        /organism="Zea mays"
        /db_xref="taxon:4577"
        /clone_lib="RNOSEQC06_SK.ab1"
        /clone_lib="Salt stressed Zea mays roots cDNA library"
        /tissue_type="Roots"
        /dev_stage="2 weeks old"
        /note="Vector: pBluescript SK+; Stressed 24 hours at 150
              mM NaCl"
  BASE COUNT
    172 a 193 c 183 g 100 t

```

Query Match	87.2%;	Score 500.6;	DB 14;	Length 648;	
Best Local Similarity	97.6%;	Pred. No. 1.8e-65;			
Matches 530;	Conservative	0;	Mismatches 9;	Indels 4;	
				Gaps	
QY	36	CACACAAGC	ACTTCGACG	TCACACGGG-CGGTGGCCACAGACACACACACACGCGTGGCGACC	94
Db	21	CACGAGAGC	ACTTCGACG	TCGCGCGGGCGGCTGTCACACAGACACACACCAAGCGTGGCGACC	80
QY	95	AATGGCTT	ACTACCAAGG	AGGTGGACTACTGTCTGGAGGAGGTGAGGTGCGTGGCCCCCGGC	15
Db	81	AATGGCTT	ACTACCAAGG	AGGTGGACTACTGTCTGGAGGAGGTGAGGTGCGTGGCCCCCGGC	14
QY	155	CGGCTTC	CGCGGCC	CCACGGCGGGCGGTCACGACGACGCTCGTCAAGGAGAGATTTCGAGGA	21
Db	141	CGGCTTC	CGCGGCC	CCACGGCGGGCGGTCGTCGAAGGAGAGATTTCGAGGA	20
QY	215	GGTCACAC	AGGTCATC	ACGGCGCGGGGCCA---ACCAACACACACCATGTGTACCAACGG	27
Db	201	GGTCACAC	AGGTCATC	ACGGCGCGGGGCCA---ACCAACACACACCATGTGTACCAACGG	26
QY	272	CGGCCAC	GGGCTTCG	TGTCGCGAGACAGGGTCCGAAGAGGACATCAACACCTGCACCGG	33
Db	261	CGGCCAC	GGGCTTCG	TGTCGCGAGACAGGGTCCGAAGAGGACATCAACACCTGCACCGG	32
QY	332	CGAGTCC	ACAGCGC	AGGAGAGCTTCTCGCCAGGGCTAACTGAGCGCGCGCGCGGC	39
Db	321	CGAGTCC	ACAGCGC	AGGAGAGCTTCTCGCCAGGGCTAACTGAGCGCGCGCGCGGC	38
QY	392	GGCATCC	ACGGCCG	TTGTTGGCTTCCGTCGGTGCCTTATGATGTCGTGGTTGACTGGTT	45
Db	381	GGCATCC	ACGGCCG	TTGTTGGCTTCCGTCGGTGCCTTATGATGTCGTGGTTGACTGGTT	44
QY	452	GTTCAGG	GTATCCT	ACTTGGCTATCTGTACGTGTGACGCACCTCAGCTCCTGTACGAAATTAC	51
Db	441	GTCAGG	GTATCCT	ACTTGGCTATCTGTACGTGTGACGCACCTCAGCTCCTGTACGAAATTAC	50
QY	512	GACAAAT	AGTCGTG	ACCTGAATAAACTTCTTCGTAATACTTAAAAA	57
Db	501	GACAAAT	AGTCGTG	ACCTGAATAAACTTCTTCGTAATACTTAAAAA	56
QY	572	AAA	574		
Db	561	AAA	563		

RESULT 8	BQ619318	648 bp	linear	EST 27-JUN-2000
LOCUS	RNOSEQD01_SK.abi	Salt stressed	zea	mays roots cDNA library
DEFINITION	mays cDNA clone RNOSEQD01_SK.abi	similar to	No homology,	mRNA
		sequence.		
ACCESSION	BQ619318			
VERSION	BQ619318.1	GI:21621312		
KEYWORDS	EST.			
SOURCE	zea	mays.		


```

|||||
501 GACAATAGCTCGTACCTGATTAACCTCTTCGTAATACATAAAAAAAAAAAAAA 560
QY 572 AAA 574
Db 561 AAA 563

RESULT 10
BO619383
LOCUS
DEFINITION
BO619383
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
FEATURES
    source
        1..648
            /organism="Zea mays"
            /db_xref="taxon:4577"
            /clone="RNOSEQ7B02_SK.ab1"
            /clone_lib="Salt stressed Zea mays roots cDNA library"
            /tissue_type="Roots"
            /dev_stage="2 weeks old"
            /note="Vector: pBluescript SK+; Stressed 24 hours at 150
mm NaCl."
BASE COUNT    172 a   193 c   183 g   100 t

Query Match      87.2%; Score 500.6; DB 14; Length 648;
Best Local Similarity 97.6%; Pred. No. 1.8e-65;
Matches 530; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

QY 36 CACACAGCATTTCGACGTCACACGGG-CGCTGGCAGACACACACCAAGCGTCGGCACC 94
Db 21 CACGAGAGCACTTCGAGTCGCGACGGCGCGTCGACACACACACCAAGCGTCGGCACC 80

QY 95 AATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCGCCGC 154
Db 81 AATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCGCCGC 140

QY 155 CGGCTTCGGCGCCGACGCGCGCGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTTCGAGGA 214
Db 141 CGGCTTCGGCGCCGACGCGCGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTTCGAGGA 200

QY 215 GGTTCGACACGTCCTCAGCGCGCGCGCCCA---ACCACCACCAACCATGTCACACGG 271
Db 201 GGTTCGACACGTCCTCAGCGCGCGCGCCCA---ACCACCACCAACCATGTCACACGG 260

QY 272 CGGCCACGGCTTCGTGGTGGCGGAGACAGGGTTCGAAGAGGACATCAACACCTGCACCGG 331
Db 261 CGGCCACGGCTTCGTGGTGGCGGAGACAGGGTTCGAAGAGGACATCAACACCTGCACCGG 320

QY 332 CGAGGTCCACGACGCGGAGGAGAGTTCCTCGCCAGGGGCTAACTAGCCCGCGCGCGCC 391
Db 321 CGAGGTCCACGACGCGGAGGAGAGTTCCTCGCCAGGGGCTAACTAGCCCGCGCGCGCC 380

```

```

QY 392 GGCATCCACGCCGCTTGGTCTGCTGCGTGCCTTATGTATGTCTGCTGGTTGACTGGTT 451
Db 381 GGCATCCACGCCGCTTGGTCTGCTGCGTGCCTTATGTATGTCTGCTGGTTGACTGGTT 440

QY 452 GTTCAGGTCATCGTACTTGGCTATCGTACGTGCACGCACCTCCTGTACGAATTAC 511
Db 441 GTGCAGGTCATCGTACTTGGCTATCGTACGTGCACGCACCTCCTGTACGAATTAC 500

QY 512 GACAATAGCTCGTACCTGAATAAACTTCTCGTAATACTAAAAA 571
Db 501 GACAATAGCTCGTACCTGAATAAACTTCTCGTAATACTAAAAA 560

QY 572 AAA 574
Db 561 AAA 563

RESULT 11
BO619390
LOCUS
DEFINITION
BO619390
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
FEATURES
    source
        1..648
            /organism="Zea mays"
            /db_xref="taxon:4577"
            /clone="RNOSEQ7B09_SK.ab1"
            /clone_lib="Salt stressed Zea mays roots cDNA library"
            /tissue_type="Roots"
            /dev_stage="2 weeks old"
            /note="Vector: pBluescript SK+; Stressed 24 hours at 150
mm NaCl."
BASE COUNT    172 a   193 c   183 g   100 t

Query Match      87.2%; Score 500.6; DB 14; Length 648;
Best Local Similarity 97.6%; Pred. No. 1.8e-65;
Matches 530; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

QY 36 CACACAGCATTTCGACGTCACACGGG-CGCTGGCAGACACACACCAAGCGTCGGCACC 94
Db 21 CACGAGAGCACTTCGACGTCGCGACGGCGCGTCGACACACACACCAAGCGTCGGCACC 80

QY 95 AATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCGCCGC 154
Db 81 AATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCGCCGC 140

QY 155 CGGCTTCGGCGCCGACGCGCGCGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTTCGAGGA 214
Db 141 CGGCTTCGGCGCCGACGCGCGCTCCAGCAGCAGTCTCGTCAAGGAGAGTTTCGAGGA 200

QY 215 GGTTCGACACGTCCTCAGCGCGCGCGCCCA---ACCACCACCAACCATGTCACACGG 271
Db 201 GGTTCGACACGTCCTCAGCGCGCGCGCCCA---ACCACCACCAACCATGTCACACGG 260

```

272 CGGCCACGGCTTCGTGGTGGCGGAGACACAGGTCAGAGGACATCAACCTGCACCGG 331
 261 CGGCCACGGCTTCGTGGTGGCGGAGACACAGGTCAGAGGACATCAACCTGCACCGG 320
 332 CGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCCGCGCGGCC 391
 321 CGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCCGCGCGGCC 380
 392 GGCATCCAGCGCCGCTTCGTGGTGGCGGAGACATCAACCTGCACCGG 451
 381 GGCATCCAGCGCCGCTTCGTGGTGGCGGAGACATCAACCTGCACCGG 440
 452 GTTCAGGTCATCGTACTTGGCTATCGTACGTCAGCGACCTCAGCTCTGTACGAATTAC 511
 441 GTTCAGGTCATCGTACTTGGCTATCGTACGTCAGCGACCTCAGCTCTGTACGAATTAC 500
 512 GACATAAGCTCGTACCTGAATAAAGCTTCCTCGTATCTTCTGATCTTCTGATCTTCTG 571
 501 GACATAAGCTCGTACCTGAATAAAGCTTCCTCGTATCTTCTGATCTTCTGATCTTCTG 560
 572 AAA 574
 561 AAA 563

RESULT 12
 BG840383
 LOCUS MEST12-H11.T7-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA
 DEFINITION sequence.
 ACCESSION BG840383
 VERSION BG840383.2 GI:14242676
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 550)
 Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
 Expressed Sequence Tags from B73 Maize Seedlings and Silks
 Unpublished (2001)
 On May 25, 2001 this sequence version replaced gi:14206705.
 Contact: Patrick S. Schnable
 Schnable Laboratory
 Iowa State University
 G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
 Tel: 515-294-0975
 Fax: 515-294-2299
 Email: schnable@iastate.edu
 PCR Primers
 FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
 BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
 Seq primer: primer T7-1 (AA TAC GAC TCA CTA TAG).
 Location/Qualifiers
 1. .550
 /organism="Zea mays"
 /lab_host="DH10B"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="MEST12-H11"
 /clone_lib="ISUM4-TN"
 /tissue_type="Seedling and silk"
 /note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
 ds-cDNA molecules were generated as follows. First-strand
 cDNA was prepared from oligo-dT selected mRNA by priming
 with a NotI oligo-dT primer (5'
 AACTGAGATTCGCGCGCGAGATTTTCTTTTCTTTT). The
 resulting DNA:RNA hybrid was treated with RNase H and used
 as a template for DNA PolI-catalyzed second strand
 synthesis. After the addition of EcoRI adaptors, the
 ds-cDNAs were digested with NotI and size-selected. The

FEATURES
 source

1. .550
 /organism="Zea mays"
 /lab_host="DH10B"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="MEST12-H11"
 /clone_lib="ISUM4-TN"
 /tissue_type="Seedling and silk"
 /note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
 ds-cDNA molecules were generated as follows. First-strand
 cDNA was prepared from oligo-dT selected mRNA by priming
 with a NotI oligo-dT primer (5'
 AACTGAGATTCGCGCGCGAGATTTTCTTTTCTTTT). The
 resulting DNA:RNA hybrid was treated with RNase H and used
 as a template for DNA PolI-catalyzed second strand
 synthesis. After the addition of EcoRI adaptors, the
 ds-cDNAs were digested with NotI and size-selected. The

resulting molecules were directionally cloned into the
 EcoRI and NotI sites of the pT73PAC vector. The library
 then went through one round of normalization to cor value
 of 5 based on the methods of Marcelo Bento Soares (Genome
 Research 6: 791-806, 1996)."
 Research 6: 791-806, 1996)."
 124 a 174 c 159 g 92 t 1 others

Query Match 86.2%; Score 494.6; DB 12; Length 550;
 Best Local Similarity 97.6%; Pred. No. 1.5e-64;
 Matches 524; Conservative 0; Mismatches 9; Indels 4; Gaps 2;
 36 CACACAGCCTTCGACGTCACACGG-CGCGTCGCGACACACACCAAGCGTGGCACC 94
 14 CCCAAGACCTTCGACGTCGACGCGCGTCGACACACACCAAGCGTGGCACC 73
 95 AATGCTTACTACACAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGCGCCGGC 154
 74 AATGCTTACTACACAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGCGCCGGC 133
 155 CGGCTTCGCGCGCCACGCGCGGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGA 214
 134 CGGCTTCGCGCGCCACGCGCGGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGA 193
 215 GGTGACACAGGTCTCAGCGCGCGGCCA---ACCACACACACACCATGTCACACGG 271
 194 GGTGACACAGGTCTCAGCGCGCGGCCA---ACCACACACACACCATGTCACACGG 253
 272 CGGCCACGGCTTCGTGGTGGCGGAGACACAGGTCGAAAGAGACATCAACCTGCACCGG 331
 254 CGGCCACGGCTTCGTGGTGGCGGAGACACAGGTCGAAAGAGACATCAACCTGCACCGG 313
 332 CGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCCGCGCGGCC 391
 314 CGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCCGCGCGGCC 373
 392 GGCATCCAGCGCCGCTTCGTGGTGGCGGAGAGCTTCCTCGCCAGGCTAACTGAGCGGCC 451
 374 GGCATCCAGCGCCGCTTCGTGGTGGCGGAGAGCTTCCTCGCCAGGCTAACTGAGCGGCC 433
 452 GTTCAGGTCATCGTACTTGGCTATCGTACGTCAGCGACCTCAGCTCTGTACGAATTAC 511
 434 GTTCAGGTCATCGTACTTGGCTATCGTACGTCAGCGACCTCAGCTCTGTACGAATTAC 493
 512 GACATAAGCTCGTACCTGAATAAAGCTTCCTCGTATCTTCTGATCTTCTGATCTTCTG 568
 494 GACATAAGCTCGTACCTGAATAAAGCTTCCTCGTATCTTCTGATCTTCTGATCTTCTG 550

RESULT 13
 AI855425/c

LOCUS 603016F02.xl 603 - stressed root cDNA library from Wang/Bohnert lab
 DEFINITION Zea mays cDNA, mRNA sequence.

ACCESSION AI855425
 VERSION AI855425.1 GI:5499558
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 539)

REFERENCE
 AUTHORS

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221

```

Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.
Location/Qualifiers
1. .539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from
Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; vector: pBluescriptII SK(+)"
Seedling stressed root cDNA library from Wang/Bohnert lab"
91 a 157 c 172 g 119 t
BASE COUNT
ORIGIN

```

Query Match	86.1%	Score 494.4	DB 9	Length 539
Best Local Similarity	97.2%	Pred. No. 1.6e-64		
Matches 525	Conservative 0	Mismatches 11	Indels 4	Gaps 2
QY	34	TCACACAAGCATTTCGACGTCACAGCGGCGCTCGGCACACACACACCAAGCGTCGGCAC	93	
DB	539	TCACACAAGCATTTCGACGTCGACGCGGCGCTGTCACACACACACCAAGCGTCGGCAC	480	
QY	94	CAATGCGTTACTACCAAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCCGTCGGTCCCGCGG	153	
DB	479	CAATGCGTTACTACCAAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCCGTCGGTCCCGCGG	420	
QY	154	CGGGCTTCGGCGCCACAGCGGCGGCGTCCAGACGACCTGCTCAAGGAGAACTTCGAGG	213	
DB	419	CGGGCTTCGGCGCCACAGGAGCGGCGTCCAGACGACCTGCTCAAGGAGAACTTCGAGG	360	
QY	214	AGGTCGACACGGTCTCACCAGCGCGGCGCA - - - ACCACCAACCAACCACCATGCTCACCAG	270	
DB	359	AGGTCGACACGGTCTCACCAGCGCGGCGCA - - - ACCACCAACCAACCACCATGCTCACCAG	300	
QY	271	GGCGCCACGGCTTCGTTGGTGC CGGAGACCAAGGTCGAAGAGGACATCAACACCTGCACCG	330	
DB	299	GGCGCCACGGCTTCGTTGGTGC CGGAGACCAAGGTCGAAGAGGACATCAACACCTGCACCG	240	
QY	331	GGAGGTCACAGCGCCAGGAGAGGTTCTCTGCCAGGGCTAACTGAGCGCGCCGCGGC	390	
DB	239	GGAGGTCACAGCGCCAGGAGAGGTTCTCTGCCAGGGCTAACTGAGCGCGCCGCGGC	180	
QY	391	CGGCATCCACGCCCGTTCGTGCTTGCGTCGTCGCTTATGTATGTCGTGGTGGTGGT	450	
DB	179	CGGCATCCACGCCCGTTCGTGCTTGCGTCGTCGCTTATGTATGTCGTGGTGGTGGT	120	
QY	451	TGTTCCAGGTCATCGTACTGTCATCGTACGTCACGCACTCAGCTCCTGTACGAATTA	510	
DB	119	TGTGAGGGTCTACTGCTTGGCTATCGTACGTCACGCACTCAGCTCCTGTACGAATTA	61	
QY	511	CGACAATAGCTCGTGACCTGAATAAACTCTTCGTAATACATAAAAAA	570	
DB	60	CGACAATAGCTCGTGACCTGAATAAACTCTTCGTAATACATAAAAAA	1	

RESULT 14	AW787732	LOCUS	AW787732	500 bp	mrna	linear	EST 16-MAY-2000
DEFINITION	(SK) Zea mays CDNA, mRNA sequence.						
ACCESSION	AW787732						
VERSION	AW787732.1						
KEYWORDS	EST.						
SOURCE	Zea mays.						
ORGANISM	Zea mays						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.						
REFERENCE	1 (bases 1 to 500)						
AUTHORS	Walbot,V.						

```

TITLE
JOURNAL
COMMENT
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. 500
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site:1: ECORI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT
97 a 159 c 156 g 88 t

```

Query Match	84.6%	Score 485.8	DB 10	Length 500
Best Local Similarity	99.6%	Prod. No. 3.2e-63		
Matches 487	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 62	GCCTCGGCACAGACACACACCAAGCGTCGGGCACCAATGGCTTACTTACCAGAGGTGGACTA	121		
Db 1	GCCTCGGCACAGACACACCAAGCGTCGGGCACCAATGGCTTACTTACCAGAGGTGGACTA	60		
QY 122	CTGCTCGGAGGAGGTGAGGTTCGGTGGCCCCCGCGCGGCTTCGGCGCGCCACGGCGCGCGCGT	181		
Db 61	CTGCTCGGAGGAGGTGAGGTTCGGTGGCCCCCGCGCGGCTTCGGCGCGCCACGGCGCGCGCGT	130		
QY 182	CCAGCAGCAGCTCGTCAAGGAGAAAGTTCGAGGAGGTTCGACACGGTCTCAGCGCCCGCGCGC	241		
Db 121	CCAGCAGCAGCTCGTCAAGGAGAAAGTTCGAGGAGGTTCGACACGGTCTCAGCGCCCGCGCGC	241		
QY 242	CAACCAACCACACACCATGGTCCACACGCGCGCCACGGCTTCGGTGGCGCCAGACCAG	301		
Db 181	CAACCAACCACACCAACCTGGTCCACGCGCGCCACGGCTTCGGTGGCGCCAGACCAG	240		
QY 302	GGTGGAGAGGAGCATCAACACCTGCACCGCGGAGGTCCACGAGCGGAGGAGAGCTTCCT	361		
Db 241	GGTGGAGAGGAGCATCAACACCTGCACCGCGGAGTTCACGAGCGGAGGAGAGCTTCCT	300		
QY 362	CGCAGGGGTAACTGAGCGCGCGCGCGCGCGGCATCCACGCGCGTTCGTTGCTGCGG	421		
Db 301	CGCAGGGGTAACTGAGCGCGCGCGCGCGGCATCCACGCGCGTTCGTTGCTGCGG	360		
QY 422	TGCGTTATGTATCTCTGTTGACCTGGTGTTCAGGGTCATCGTACTTGGCTATCGTAC	481		
Db 361	TGCGTTATGTATCTCTGTTGACCTGGTGTTCAGGGTCATCGTACTTGGCTATCGTAC	420		
QY 482	GTGCAGGCACCTCAGCTCCCTGTACGAATTACGAAATTAAGCTTCGTGACCTGAATAAATTT	541		
Db 421	GTGCAGGCACCTCAGCTCCCTGTACGAATTACGAAATTAAGCTTCGTGACCTGAATAAATTT	541		
QY 542	CTTCGGTAAT	550		
Db 481	CTTCGGTAAT	489		

RESULT 15

AW787314

LOCUS AW787314 496 bp mRNA linear EST 16-MAY-2000
DEFINITION 945002506.X1 945 - Mixed adult tissues from Walbot lab, same as 707
(SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW787314
VERSION AW787314.1 GI:7844111
KEYWORDS EST.
SOURCE Zea mays.

ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 496)

REFERENCE Walbot.V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. .496
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 97 a 156 c 153 g 90 t
ORIGIN

Query Match 81.1%; Score 465.6; DB 10; Length 496;
Best Local Similarity 99.0%; Pred. No. 3.2e-60;
Matches 479; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY	68	CGCACACACACACACGCTCGGCACCAATGGCTTACTACGAGGAGTGACTACTGCTC	127
DB	1	CGCACACACACACACGCTCGGCACCAATGGCTTACTACGAGGAGTGACTACTGCTC	60
QY	128	GGAGGAGTGAGTGCCTGCGCCCGCGGCTTCGGCCGCCACGCGCGGCGGTCCAGCA	187
DB	61	GGAGGAGTGAGTGCCTGCGCCCGCGGCTTCGGCCGCCACGCGCGGCGGTCCAGCA	120
QY	188	GCACGTCGTCAGGAGAGTTCAGGAGGTCGACAGGTCACGGTCTCAGCGCGCGGCGCA	246
DB	121	GCACGTCGTCAGGAGAGTTCAGGAGGTCGACAGGTCACGGTCTCAGCGCGCGGCGCA	180
QY	247	ACCACACACACACACGCTCGGCACCGCGGCGGCTTCGGTTCGCGGAGACACAGGTCG	306
DB	181	ATCACACACACACACGCTCGGCACCGCGGCGGCTTCGGTTCGCGGAGACACAGGTCG	240
QY	307	AAGAGGACATCAACACCTGCACCGGAGAGTTCACGAGCGCGGAGAGAGCTTCCTCGCCA	366
DB	241	AAGAGGACATCAACACCTGCACCGGAGAGTTCACGAGCGCGGAGAGAGCTTCCTCGCCA	300
QY	367	GGGCTAACTAGCGCGCGCGCGGCGGCGGCTTCGGTTCGCTGCTGCTGCTGCTGCT	426
DB	301	GGGCTAACTAGCGCGCGCGCGGCGGCGGCTTCGGTTCGCTGCTGCTGCTGCTGCT	360
QY	427	TATGTATGTCTGTGTTGACTGTTGTTACAGGTCATCGTACTGTTGCTATCGTACGTGCA	486

Mon Jun 16 14:55:59 2003

```

ID Q9SM40 PRELIMINARY; PRT; 95 AA.
AC Q9SM40;
DT 01-JUN-2002 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-WAR-2001 (TREMBlrel. 16, Last annotation update)
DE Putative glycine-rich protein.
OS Sporobolus stapfianus (Resurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade;
OC Chloridoideae; Eragrostideae; Sporobolus.
NCBI_TaxID=56623;
[1]
SEQUENCE FROM N.A.
RA Neale A.D., Blomstedt C.K., Bronson P., Le T.N., Guthridge K.,
RA Evans J., Gaff D.F., Hamill J.D.;
RA "The isolation of lowly-transcribed genes which are induced during
RT desiccation of the resurrection grass Sporobolus stapfianus."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ242802; CAB61838.1;
DR EMBL: AJ242802; CAB61838.1;
SQ SEQUENCE 95 AA; 10668 MW; D756DCE2B68DD85B CRC64;

Query Match 34.7%; Score 177; DB 10; Length 95;
Best Local Similarity 57.7%; Pred. No. 1.5e-12;
Matches 45; Conservative 8; Mismatches 15; Indels 10; Gaps 6;

Qy 1 MAYTQE-VDYCSEVRVAPAGFGRH-GGGVQVHVKEFEVDVTSRAGANHHHHGHH 58
Db 1 MAHFKENVDY--EVTSMKPGRGHGGGVQVVKTFEEVQVTPGRS-GHHGHR 57

Qy 59 G--CHG---FVRETRVE 71
Db 58 GNOGHSGHFQARETFE 75

RESULT 3
Q8S0B2 PRELIMINARY; PRT; 102 AA.
AC Q8S0B2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE P0470A12.4 protein.
GN P0470A12.4
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
NCBI_TaxID=39947;
[1]
SEQUENCE FROM N.A.
RA STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0470A12."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003436; BAB0279.1;
DR EMBL: AP003436; BAB0279.1;
SQ SEQUENCE 102 AA; 11214 MW; 6FF1266B1CDE7768 CRC64;

Query Match 27.6%; Score 141; DB 10; Length 102;
Best Local Similarity 45.8%; Pred. No. 2e-08;
Matches 27; Conservative 8; Mismatches 18; Indels 6; Gaps 1;

Qy 3 YVEVDYCSSEVRVAPAGFGRHGGGVQVHVKEFEVDVTSRAGANHHHHGHHG 61
Db 4 YVEVDYCSSEVRVAPAGFGRHGGGVQVHVKEFEVDVTSRAGANHHHHGHHG 56

RESULT 4
Q93ZA6 PRELIMINARY; PRT; 196 AA.
ID Q93ZA6;
AC Q93ZA6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

```

SQ      SEQUENCE      196 AA;  21362 MW;  0023D5571345C6A0 CRC64;

Query Match      18.1%;  Score 92.5;  DB 10;  Length 196;
Best Local Similarity 33.8%;  Pred. No. 0.013;
Matches 24;  Conservative 12;  Mismatches 24;  Indels 11;  Gaps 5

QY      21  GFGRHGGG-VQOHHVKEKFE- - - - -VDTVSRA- - - - -NHHHHHGH- HGHGFGFVRET- - - R 69
      | | | | | : : : | | | | | | | | | | | | | | | | | | |
Db      99  GGGVNGGDMSAAVIRKAAEVGAQVDLEAAGAGGNRRHHHHHGHQGRNHDYVDNHSYR 158

QY      70  VEEDINTCTGE 80
      : : | : | :
Db      159  INDDLMCECSK 169

```

RESULT	6
O24643	
ID	O24643
AC	O24643; PRELIMINARY; PRT; 259 AA.
DT	01-JAN-1998 ('TREMBlrel. 05, Created)
DT	01-JAN-1998 ('TREMBlrel. 05, Last sequence update)
DT	01-JUN-2002 ('TREMBlrel. 21, Last annotation update)
DE	TNY-like protein (Fragment).
OS	Arabidopsis thaliana (Mouse-ear cross).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC	Spermatocytas; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX	eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC	NCBI_TaxId=3702;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=COLUMBIA;
RA	Terryn N.;
RL	Submitted (NOV-1997) to the ENBL/GenBank/DDBJ databases
RN	[2]

RX	MEDLINE-97338147; PubMed-9192694;
RA	Okumuro J.K., Caster B., Villarroel R., van Montagu M., Jofuku K.D.;
RT	"The AP2 domain of APEFAL2 defines a large new family of DNA binding
RT	proteins in Arabidopsis.";
RL	Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).
DR	ENBL; AJ002598; CAA05630.1; -;
DR	EMBL; AF003103; AAC49776.1; -;
DR	HSP; O80337; ZGCC;
DR	InterPro; IPR001471; TF_ERF.
DR	Pfam; PF00847; AP2-domain; 1.
DR	PRINTS; PR00367; ETHRSELEMNT.
DR	PRODOM; PD001423; TF_AP2; 1.
DR	SMART; SMO0380; AP2; 1.
FT	NON_TER 1
FT	CHAIN 64 >259
FT	NON_TER 259
FT	SEQUENCE 259 AA; 28615 MW; 0464949DB6C619DB CRC64:
SQ	

```

Query Match      18.1%; Score 92.5; DB 10; Length 259;
Best Local Similarity 33.8%; Pred. No. 0.018;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY      21  GFGRRGGG-VQOHVVKKEFEE----VDTVSRAGA---NHHHHHHGH-HGCHGFVYVRET--R 69
          | | | | | : : | | | : | | | | | : | | | : | | | : |
Db      162  GGVNGVGGDMSAAVIRKKAEGVQYDALEAGAGGNNRRHHHHHHQHQRNHDYDVNHSYR 221

QY      70  VEEDINTCTGE 80
          : : | : | :
Db      222  INDDLMCCSSK 232
          : : | : | :

```

RESULT 7
Q9P543 PRELIMINARY: PRT; 1245 AA.
ID Q9P543
AC Q9P543;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE Related to multiform cyclin-dependent kinase PHO85.
GN B24H17.10.
OS Neurospora crassa,
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
[1]
RP Submitted FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nykattula G., Meves H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL: AL356815; CAB92623.1; -.
DR HSSP: P42773; LIIB...
DR InterPro: IPR002110; ANK.
DR InterPro: IPR004129; GDPD.
DR InterPro: IPR004331; SPX.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF03009; GDPD; 1.
DR Pfam: PF03105; SPX; 1.
DR SMART: SM00248; ANK; 3.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS0297; ANK_REPEAT_REGION; 1.
DR ANK repeat; Cyclin; Kinase; Repeat.
SQ SEQUENCE 1245 AA; 137798 MW; 16BD0F6A04596A9F CRC64;

```

Query Match      17.5%  Score 89;  DB 3;  Length 1245;
Best Local Similarity 72.2%  Pred. No. 0.29;
Matches 13;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

QY      45  SRAGANHHHHHHGGHGGH 62
          ||:|:||||| ||| ||
Db       901  SRSGTTHHHHHHHHHGGH 918

```

RESULT 8

Q850B4	PRELIMINARY;	PRT;	408 AA.
ID	Q850B4		
Q850B4;			
AC	Q850B4;		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	P0470A12.2 protein.		
DE	P0470A12.2		
GN	P0470A12.2		
OS	Oryza sativa [japonica cultivar-group].		
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoidae; Oryzeae; Oryza.		
NCBI_TaxID=39947;			
ON	NCBI_TaxID=39947;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC		
RT	clone:P0470A12.";		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AP0003436; BAB90277.1; --		
DR	SEQUENCE 408 AA; 43466 MW; F760297877F6D69C CRC64;		
SQ	SEQUENCE		

Query Match	17.0%;	Score 86.5;	DB 10;	Length 408;
Best Local Similarity	26.9%;	Pred. No. 0.15;		
Matches	25;	Conservative 10;	Mismatches 15;	Indels 43; Gaps 5;
Qy	12	EEVRSVAPAGFRH-----GGGVQHVYKSKFEVDTVSRAGA-----NHH-----	52	
Db	199	EECDAAAGGYGRHHGAGGVAKQHATYKQHQALIEGVNCGGAGYNYNRHQAVAVAGGQ	258	
Qy	53	HHH-----GGH-----GGH 61		
Db	259	HHYGATAAAYGNASNKQHTAAAGHHSSGGH	291	


```

RESULT 9
Q9UAY0 PRELIMINARY; PRT; 471 AA.
Q9UAY0
AC QUAYO;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE W03G1.5 protein..
DE W03G1.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RA Pauley A., Scheet P., Harper M.;
RT "The sequence of C. elegans cosmid W03G1.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF125964; AAD14753.1;
SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A985B CRC64;
Query Match 17.0%; Score 86.5; DB 5; Length 471;
Best Local Similarity 34.8%; Pred. No. 0.17; Indels 25; Gaps 4;
Matches 24; Conservative 4; Mismatches 16;
QY 15 RSVAPAGRGRRGG-----GVOQHVVVKEFEVDVVSAGANHHHHHG----- 56
DB 372 RSHSPRGHG-HGRRGHPCHPCGRHHGPPHH-----HHHGDGRSPSRGHHHHHHHGGCRPF 426
QY 57 --HHGHHGF 63
DB 427 PPHGHHHF 435
RESULT 10
Q9Y2L9 PRELIMINARY; PRT; 793 AA.
Q9Y2L9
AC Q9Y2L9;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE KIAA1016 protein (Fragment).
DE KIAA1016
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]

```


Db 377 DYISEFV-SLLPKSIRR-----VAEPIPEEVQKVL EAKAGDDHHHHGHGHAHAGY 428

RESULT 12
Q9FZ33
ID Q9FZ33 PRELIMINARY; PRT: 473 AA.
AC Q9FZ33;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE T24C10.10 protein (Hypothetical 52.4 kDa protein).
T24C10.10 OR F14C21.51 OR ATIG54990.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Embryophyta; Tracheophyta;
OC eucosids ii; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC NCBI_TaxID=3702; Brassicales; Brassicaceae; Arabidopsi
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Alfafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chlou J., Choi E., Gonzalez S.,
RA Howing B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharasy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Eckert J., Theologis A., Davis R.W.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Eckert J.R., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA White O., Alonso J., Altabi H., Chou Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Chan A., Chao Q., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hulzar L.,
RA Hunter J.B., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.S., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Malit R., Marziali A.,
RA Militcher J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
RA Davis R.W., Eckert J.R., Theologis A.;
RT "Full length cDNA of gene AtG54990 (GI:15221965).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC064840; AAG00878.1; -
DR EMBL; AC069144; AAG51115.1; -
DR EMBL; AY074334; AAL67030.1; -
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 52419 MW; D09124A11565BB23 CRC64;

```

Query Match      16.5%  Score 84;  DB 10;  Length 473;
Best Local Similarity 36.8%  Pred. No. 0.34;
Matches 21;  Conservative 10;  Mismatches 18;  Indels 8;  Gaps 3;

QY      8  DYCSEVRSVAPAGCGRHGGVGOHVVKEFEEVDTVSRAGANHHHHHHG---HGHH 61
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

Db 409 DYISEFV-SLLPKSIRR-----VAEPIPEEVQKVLEAKAGDDHDHIIHGHGHAAHAGY 460

RESULT	13
Q9W3D2	
ID	PRELIMINARY; PRT; 1561 AA.
AC	Q9W3D2;
DT	01-MAY-2000 (TRENBLrel. 13, Created)
DC	01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT	01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE	CRAG protein.
GN	CRAG OR CG12737.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
QC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BERKELEY.
RX	MEDLINE-20196006; PubMed-107311132;
RA	Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D., Ananadis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.C., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D., Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris D., Haussler K.A., Howland T.J., Hernandez J.R., Houck J., Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.G., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris D., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusstern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Spidan-Klamas I., Simpson M., Skupski M.P., Smith T., Spiers R., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; The genome sequence of <i>Drosophila melanogaster</i> .

Science 287:2185-2195 (2000).
ENBL; AE003444; AAF46397.1; FlyBase; FBgn0025864; Crag. InterPro: IPRO05112; dENN. InterPro: IPRO01194; DENN. InterPro: IPRO05113; uENN. Pfam; PF03455; dENN; 1. Pfam; PF02141; DENN; 1. Pfam; PF03456; uENN; 1. Q SEQUENCE 1561 AA; 174516 MW; 10BFDF38A3DD4FC4E CRC64;

Query Match 16.5%; Score 84; DB 5; Length 1561;
Best Local Similarity 23.9%; Pred. No. 1.4;
Matches 28; Conservative 17; Mismatches 30; Indels 42; Gaps 5;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:48:33 ; Search time 25 Seconds
(without alignments)
357.620 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQVEDYCEEVRSVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	92.5	18.1	196	2	G85435
2	92.5	18.1	259	2	T52619
3	92	18.0	414	2	A48273
4	89	17.5	1245	2	T49815
5	86.5	17.0	471	2	T33997
6	84	16.5	473	2	D96591
7	82.5	16.2	735	2	T45059
8	82	16.1	173	2	T51469
9	81.5	16.0	495	1	S31223
10	81	15.9	86	2	T16437
11	81	15.9	191	2	A42844
12	80	15.7	133	2	T09608
13	80	15.7	1273	2	T00338
14	79	15.5	204	2	T09592
15	79	15.5	414	2	A40350
16	77	15.1	754	2	JC4898
17	76.5	15.0	77	2	T16436
18	76.5	15.0	332	2	C83682
19	76	14.9	102	2	T30119
20	76	14.9	1560	2	T00080
21	75.5	14.8	425	1	JH0710
22	75	14.7	369	1	TVFVAF
23	75	14.7	600	2	F71434
24	75	14.7	1019	2	T00117
25	74.5	14.7	1441	2	T13717
26	74.5	14.6	410	2	T26757
27	74	14.5	83	2	T16435
28	74	14.5	.895	2	JC7089
29	74	14.5	1891	2	T13594

30 73.5 14.4 147 2 T16440
31 73.5 14.4 285 2 E84766
32 73.5 14.4 499 2 S09880
33 73.5 14.4 776 2 T02702
34 73 14.3 606 2 A41145
35 73 14.3 633 1 A26030
36 72.5 14.2 1585 2 T31611
37 72 14.1 354 2 S39406
38 71.5 14.0 268 2 A56446
39 71.5 14.0 499 2 C75251
40 71.5 14.0 699 2 A54660
41 71 13.9 338 2 A82890
42 71 13.9 361 2 F87286
43 71 13.9 2038 2 A43742
44 71 13.9 2396 2 T13714
45 70.5 13.8 59 2 T09610

ALIGNMENTS

RESULT 1
G85435
TINY-like protein [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85435
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <STO>
A:Cross-references: GB:NC_001268; NID:97270639; PIDN:CAB80356.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g36900
A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.012; 24; Mismatches 12; Indels 11; Gaps 5;
Matches 24; Conservative 12; Indels 11; Gaps 5;

QY 21 GFGRRGGG-VQHHVYKKEE-----VDTVSRAGA---NHHHHHGH-HGGHGVVYRET--R 69
Db 99 GGGVNGGDMSSAAVIRKAAEVGAQVDALEAAGAGGNRRHHHHHHQHQRGNDYVDNHSYR 158
QY 70 VEEDINTCTGE 80
Db 159 INDDLMECSSK 169

RESULT 2
T52619
TINY-like protein [Imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C:Accession: T52619
R:Terryn, N.; Heijnen, L.; De Keyser, A.; Van Asseldonck, M.; De Clercq, R.; Verbakel
ueller, C.; Mayer, K.; Dehais, P.; Rombauts, S.; Van Montagu, M.; Rouze, P.; Vos, P.
FEBS Lett. 445, 237-245, 1999
A:Title: Evidence for an ancient chromosomal duplication in Arabidopsis thaliana by
A:Reference number: 226022; MUID:99192287; PMID:10094464
A:Accession: T52619
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-259 <TER>
A:Cross-references: EMBL:AJ002598; PIDN:CAA05630.1
C:Genetics:
A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 259;

hypothetical prote
probable AT-hook D
hypothetical prote
hypothetical prote
gamma-aminobutyric
serine/threonine-s
hypothetical prote
homeotic protein o
Ig heavy chain v r
phosphorylase hy
histidine rich cal
hypothetical prote
cation efflux fami
female sterile hom
kakapo gene protel
environmental stre

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1245 <SCH>
A;Cross-references: EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.10
A;Experimental source: BAC clone B24H17; strain OR74A
C;Genetics:
A;Gene: NCSP:B24H17.10
A;Map position: 6
A;Introns: 6/2; 1141/3

Query Match 17.5%; Score 89; DB 2; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.21;
Matches 13; Conservative 2; Mismatches 0; Gaps 0;

QY 45 SRAGANHHHHHGGHGG 62
Db 901 SRSGTHHHHHHHGGHGG 918

RESULT 5
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33997
R;Pauley, A.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid W03G1.
A;Reference number: 221454
A;Accession: T33997
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-471 <PAU>
A;Cross-references: EMBL:AF125964; PIDN:AA14753.1; GSPDB:GN00022; CESP:W03G1.5
A;Experimental source: strain Bristol N2; clone W03G1
C;Genetics:
A;Gene: CESP:W03G1.5
A;Map position: 4

Query Match 17.0%; Score 86.5; DB 2; Length 471;
Best Local Similarity 34.8%; Pred. No. 0.14;
Matches 24; Conservative 4; Mismatches 16; Indels 25; Gaps 4;

QY 15 RSVAPAGGFRGGG-----GQOVVKEFEVDVTSRAGANHHHHG----- 56
Db 372 RSHSPRGHG-HGGRHPPHPCGRHGHGPHH-----HHHDKGSPRHHHHHHHGGCRPF 426

QY 57 --HHGGHGF 63
Db 427 PPHGGHHHF 435

RESULT 6
D96591
hypothetical protein T24C10.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana [mouse-ear cross]
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96591
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>

Best Local Similarity 33.8%; Pred. No. 0.016;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY 21 GFRHGGG-VQOVVKEFE-----VDVTSRAGA---NHHHHHGG-HGGHGFVVRER--R 69
Db 162 GGVGVGGGMSAAVIRKAAEAGVQVDALEAAGAGNRHHHHHGHQGRNHDYVDNHSYR 221

QY 70 VEEDINTCTGE 80
Db 222 INDDLMECSSK 232

RESULT 3
A48273
delta/Y1/NF-EL/UCRBP transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C;Accession: A48273; A42055; A56418
R;Safrany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A;Title: Characterization of the mouse gene that encodes the delta/Y1/NF-EL/UCRBP trans
A;Reference number: A48273; MUID:93296177; PMID:8516301
A;Accession: A48273
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-414 <RES>
A;Cross-references: GB:L13968; NID:g293847; PIDN:AAA40477.1; PID:g293849
R;Flanagan, J.R.; Becker, K.G.; Ennist, D.L.; Gleason, S.L.; Driggers, P.H.; Levi, B.Z.;
Mol. Cell. Biol. 12, 38-44, 1992
A;Title: Cloning of a negative transcription factor that binds to the upstream conserved
A;Reference number: A42055; MUID:92107191; PMID:1309593
A;Accession: A42055
A;Molecule type: mRNA
A;Residues: 1-414 <FLA>
A;Cross-references: GB:M73963; NID:g202270; PIDN:AAA40522.1; PID:g202271
A;Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:74642)
R;Harharan, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A;Title: delta, a transcription factor that binds to downstream elements in several poly
A;Reference number: A56418; MUID:92052178; PMID:1946404
A;Accession: A56418
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218, 'S', 220-374, 'G', 376-414 <HAR>
A;Cross-references: GB:M74590; NID:g192940; PIDN:AAA37521.1; PID:g192941
C;Genetics:
A;Introns: 227/1; 281/2; 301/3; 354/3
C;Keywords: transcription factor; zinc finger

Query Match 18.0%; Score 92; DB 2; Length 414;
Best Local Similarity 31.0%; Pred. No. 0.031;
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps 2;

QY 21 GFRHGGG-VQOVVKEFEVDVTSRAGANHHHHHGGHGGHGFVVRERTRVEDINTCTGE 80
Db 55 GGGDHGGGGGGH-----GHAGHHHHHHHHHHHPPMIALQPLVTD-----PTQ 98

QY 81 VHERRESFLAR 91
Db 99 VHHQEVILVQ 109

RESULT 4
T49815
related to multifunctional cyclin-dependent kinase PHO85 [imported] - Neurospora crassa
N;Alternate names: protein B24H17.10
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49815
R;Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: 225022
A;Accession: T49815

A:Cross-references: GB:AE0051173; NID:g9857523; PIDN:AG00878.1; GSPDB:GN00141
C:Genetics:
A:Gene: T24C10.10
A:Map position: 1

Query Match 16.5%; Score 84; DB 2; Length 473;
Best Local Similarity 36.8%; Pred. No. 0.26;
Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;
QY 8 DYCSEEVRSVAPAGFGRHGGVQVHVKEFEVDVTSRAGANHHHHHGHG---HGHH 61
DB 409 DYISEFV-SLLPKSIRR-----VAEPIPEVQKLEAKAGDDHDDHHGHGHAHAGY 460

RESULT 7
T45059
hypothetical protein Y39B6B.gg [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45059
R:Wilson, R.; Alnsough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.; Nature 368, 32-38, 1994
A:Authors: Showkeen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonhammer, E.; S. tock, L.; Wilkinson-Sproat, J.; Wohldman, P.
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID:94150718; PMID:7906398
A:Accession: T45059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-735 <WIL>
A:Cross-references: EMBL:AL132896; NID:g6434440; PIDN:CAB60938.1; PID:g6434473
A:Experimental source: clone Y39B6B
C:Genetics:
A:Map position: 3
A:Introns: 18/1; 69/1
A:Note: Y39B6B.gg

Query Match 16.2%; Score 82.5; DB 2; Length 735;
Best Local Similarity 26.4%; Pred. No. 0.6;
Matches 19; Conservative 1; Mismatches 9; Indels 43; Gaps 2;
QY 23 GRHGGVQVHVKEFEVDVTSRAGANHHH-----HGHHGAAHHHAPHHEHGHGHHGSHGVHGHGTHH 53
DB 614 GHGAGYGAH-----HGHHGAAHHHAPHHEHGHGHHGSHGVHGHGTHH 662
QY 54 ---HHGHHGGHG 62
DB 663 SLAHGHHGGHG 674

RESULT 8
T51469
glycine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein K10A8_130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51469
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <SAT>
A:Cross-references: EMBL:AL391151
A:Experimental source: cultivar Columbia; BAC clone K10A8
C:Genetics:
A:Map position: 5
A:Introns: 97/1
A:Note: K10A8_130

Query Match 16.1%; Score 82; DB 2; Length 173;
Best Local Similarity 34.7%; Pred. No. 0.15;
Matches 17; Conservative 7; Mismatches 15; Indels 10; Gaps 2;
QY 19 PAGFGRHGGVQVHVKEFEVDVTSRAGA---NHHHHHHGHGHHG 63
DB 93 PSHSGHHGGTGATGAIAG-----GVAAGAAAHMHSHHHGHGHHGHG 135

RESULT 9
S31223
transcription factor Brn-1 - mouse
N:Alternate names: class III POU domain protein brain-1
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
C:Accession: S31223
R:Hara, Y.; Roverscalli, A.C.; Kim, Y.; Mirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A:Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A:Reference number: S31223; MUID:92228768; PMID:1565620
A:Accession: S31223
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAR>
A:Cross-references: EMBL:M88299; NID:g200444; PIDN:AAA39960.1; PID:g200445
C:Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:27-49/Region: glycine-rich
F:101-112/Region: alanine-rich
F:162-180/Region: histidine/proline-rich
F:186-201/Region: alanine-rich
F:236-247/Region: glycine-rich
F:267-291/Region: histidine/proline-rich
F:316-383/Domain: POU domain homology <POU>
F:402-458/Domain: homeobox homology <HOK>

Query Match 16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.5;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;
QY 19 PAGFGRHGGVQVHVKEFEVDVTSRAGANHHHH-----GHGHHG- 62
DB 237 PGGGGGAGGAGSLVHPLVGRDTPELAHHHHHHHHHPPHPPHHAQGGG 296
QY 63 -----FVYRETRVEEDINTCTGEVHER 85
DB 297 GPLNSHDPHSDPTSDLEQFAKQKRR 328

RESULT 10
T16437
hypothetical protein F53A9.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16437
R:Miller, N.
submitted to the EMBL Data Library, March 1995
A:Description: The sequence of C. elegans cosmid F53A9.
A:Reference number: Z18513
A:Accession: T16437
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <MIL>
A:Cross-references: EMBL:U23523; NID:g746551; PID:g746557; PIDN:AAC46561.1; CESP:F53A
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F53A9.6

Query Match 15.9%; Score 81; DB 2; Length 86;
Best Local Similarity 29.5%; Pred. No. 0.089;
Matches 23; Conservative 5; Mismatches 24; Indels 26; Gaps 3;
QY 1 MAYQYVDYCYSEEVRSVAPAGFGRHGG-----GVQHHVYKKEFEVDVTS 45

C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00338
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <NAG>
A:Cross-references: EMBL:AB011142; NID:g3043663; PIDN:BAAJ5496.1; PID:g3043664
A:Experimental source: brain; clone HH2365
C:Genetics:
C:Superfamily: human hypothetical protein KIAA0570

Query Match 15.7%; Score 80; DB 2; Length 1273;
Best Local Similarity 33.3%; Pred. No. 2;
Matches 15; Conservative 9; Mismatches 19; Indels 2; Gaps 1;

QY 48 GANHHHHGGHGGH--GFVVRTRVEDINTCTGEVHERRESFLA 90
DB 595 GPOHHHHHHHHHHHHGHWDDMLSDADDYSCSSQVSASEKNMA 639

RESULT 14
T09592
protein corA, cold- and drought-regulated - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09592
R:Laberge, S.; Castonguay, Y.; Vezina, L.P.
Plant Physiol. 101, 1411-1412, 1993
A:Title: New cold- and drought-regulated gene from Medicago sativa.
A:Reference number: Z16754; MUID:94143496; PMID:8310076
A:Accession: T09592
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-204 <LAB>
A:Cross-references: EMBL:L03708; NID:g289122; PIDN:AAA99833.1; PID:g289123
C:Genetics:
C:Gene: corA
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 15.5%; Score 79; DB 2; Length 204;
Best Local Similarity 30.4%; Pred. No. 0.37; Indels 28; Gaps 3;
Matches 24; Conservative 3; Mismatches 24; Indels 28; Gaps 3;

QY 21 GFRHGGGVQOHV-VKEFEVDV-----SRAGNH-----H 52
DB 125 GHGGHGAESVAVQTEKNEVDNDAKYGGSYNDGRGYNHGGGYNHGGGGHGGH 184

QY 53 HHHHHGGHGFVVRTRVE 71
DB 185 GGHGGHGGHGA VQTEEDNTQ 203

RESULT 15
T40350
transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 08-Oct-1999
C:Accession: A40350; S78494; S33712; A56419
R:Shi Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and
A:Reference number: A40350; MUID:92005716; PMID:1655281
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: GB:M77698; NID:g186767; PIDN:AAA59467.1; PID:g186768

C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00338
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <NAG>
A:Cross-references: EMBL:AB011142; NID:g3043663; PIDN:BAAJ5496.1; PID:g3043664
A:Experimental source: brain; clone HH2365
C:Genetics:
C:Superfamily: human hypothetical protein KIAA0570

Query Match 15.7%; Score 80; DB 2; Length 1273;
Best Local Similarity 33.3%; Pred. No. 2;
Matches 15; Conservative 9; Mismatches 19; Indels 2; Gaps 1;

QY 48 GANHHHHGGHGGH--GFVVRTRVEDINTCTGEVHERRESFLA 90
DB 595 GPOHHHHHHHHHHHHGHWDDMLSDADDYSCSSQVSASEKNMA 639

RESULT 14
T09592
protein corA, cold- and drought-regulated - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09592
R:Laberge, S.; Castonguay, Y.; Vezina, L.P.
Plant Physiol. 101, 1411-1412, 1993
A:Title: New cold- and drought-regulated gene from Medicago sativa.
A:Reference number: Z16754; MUID:94143496; PMID:8310076
A:Accession: T09592
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-204 <LAB>
A:Cross-references: EMBL:L03708; NID:g289122; PIDN:AAA99833.1; PID:g289123
C:Genetics:
C:Gene: corA
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 15.5%; Score 79; DB 2; Length 204;
Best Local Similarity 30.4%; Pred. No. 0.37; Indels 28; Gaps 3;
Matches 24; Conservative 3; Mismatches 24; Indels 28; Gaps 3;

QY 21 GFRHGGGVQOHV-VKEFEVDV-----SRAGNH-----H 52
DB 125 GHGGHGAESVAVQTEKNEVDNDAKYGGSYNDGRGYNHGGGYNHGGGGHGGH 184

QY 53 HHHHHGGHGFVVRTRVE 71
DB 185 GGHGGHGGHGA VQTEEDNTQ 203

RESULT 15
T40350
transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 08-Oct-1999
C:Accession: A40350; S78494; S33712; A56419
R:Shi Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and
A:Reference number: A40350; MUID:92005716; PMID:1655281
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: GB:M77698; NID:g186767; PIDN:AAA59467.1; PID:g186768

C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00338
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <NAG>
A:Cross-references: EMBL:AB011142; NID:g3043663; PIDN:BAAJ5496.1; PID:g3043664
A:Experimental source: brain; clone HH2365
C:Genetics:
C:Superfamily: human hypothetical protein KIAA0570

Query Match 15.7%; Score 80; DB 2; Length 1273;
Best Local Similarity 33.3%; Pred. No. 2;
Matches 15; Conservative 9; Mismatches 19; Indels 2; Gaps 1;

QY 48 GANHHHHGGHGGH--GFVVRTRVEDINTCTGEVHERRESFLA 90
DB 595 GPOHHHHHHHHHHHHGHWDDMLSDADDYSCSSQVSASEKNMA 639

RESULT 14
T09592
protein corA, cold- and drought-regulated - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09592
R:Laberge, S.; Castonguay, Y.; Vezina, L.P.
Plant Physiol. 101, 1411-1412, 1993
A:Title: New cold- and drought-regulated gene from Medicago sativa.
A:Reference number: Z16754; MUID:94143496; PMID:8310076
A:Accession: T09592
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-204 <LAB>
A:Cross-references: EMBL:L03708; NID:g289122; PIDN:AAA99833.1; PID:g289123
C:Genetics:
C:Gene: corA
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 15.5%; Score 79; DB 2; Length 204;
Best Local Similarity 30.4%; Pred. No. 0.37; Indels 28; Gaps 3;
Matches 24; Conservative 3; Mismatches 24; Indels 28; Gaps 3;

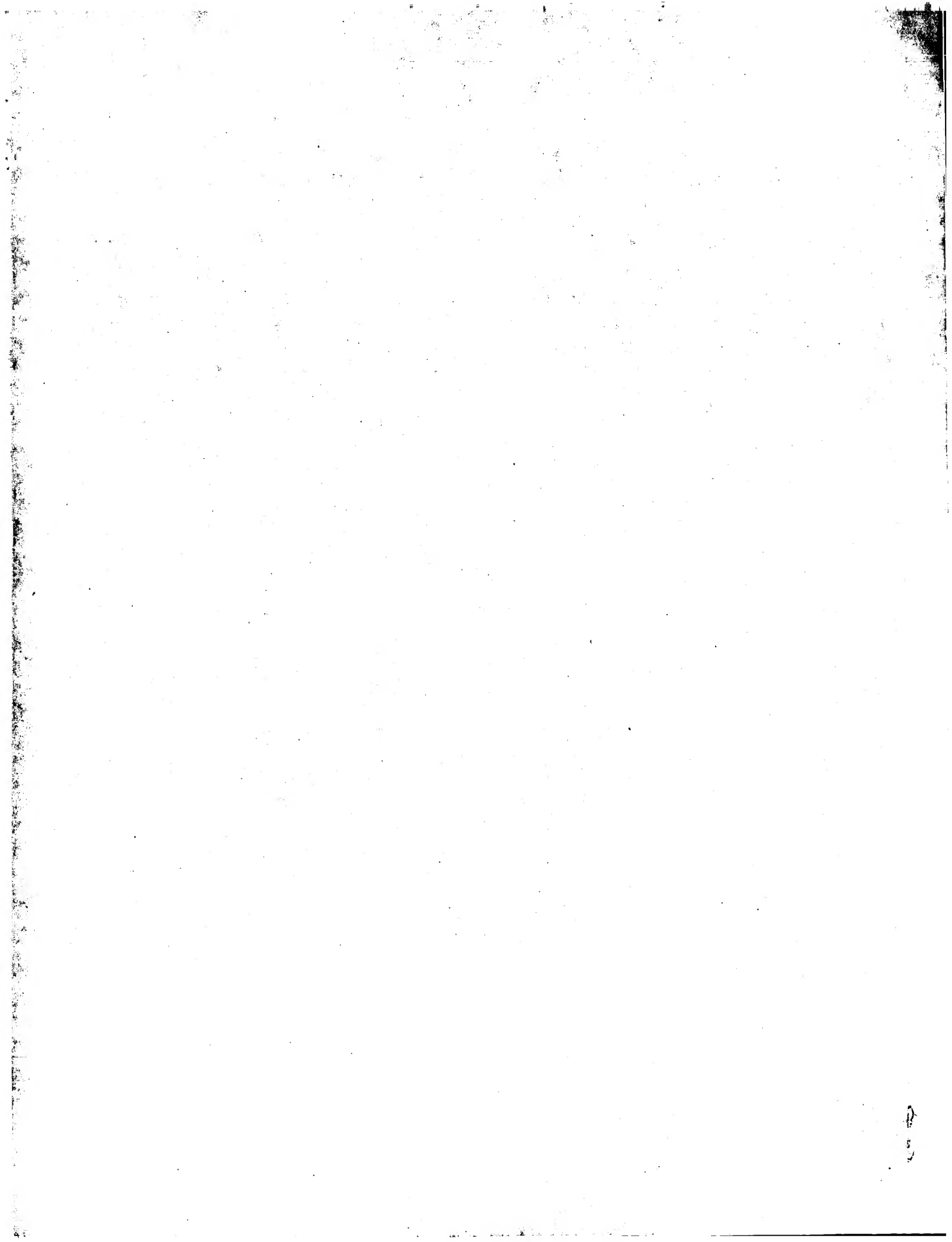
QY 21 GFRHGGGVQOHV-VKEFEVDV-----SRAGNH-----H 52
DB 125 GHGGHGAESVAVQTEKNEVDNDAKYGGSYNDGRGYNHGGGYNHGGGGHGGH 184

QY 53 HHHHHGGHGFVVRTRVE 71
DB 185 GGHGGHGGHGA VQTEEDNTQ 203

RESULT 15
T40350
transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 08-Oct-1999
C:Accession: A40350; S78494; S33712; A56419
R:Shi Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and
A:Reference number: A40350; MUID:92005716; PMID:1655281
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: GB:M77698; NID:g186767; PIDN:AAA59467.1; PID:g186768

C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
R:N

Search completed: June 6, 2003, 12:53:25
Job time : 26 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 11:41:43 ; Search time 125 Seconds
(without alignments)
1077.357 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQEVDCYSEEVSRVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10090035/runat_06062003_105507_11116/app_query.fasta_1.526
-DB=Published Applications_NA -QFMT=fastap -SURFIX=rpnp -MINMATCH=0.1
-LOOCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10090035_8CGN_1_176_@unat_06062003_105507_11116
-NCPU=6 -ICPU=3 -NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	574	9 US-10-090-035-3	Sequence 3, Appli
2	510	100.0	577	9 US-10-090-035-5	Sequence 5, Appli
3	510	100.0	676	9 US-10-090-035-1	Sequence 1, Appli
4	507	99.4	529	9 US-10-090-035-9	Sequence 9, Appli

5	499.5	97.9	524	9	US-10-090-035-17	Sequence 17, Appli
6	499.5	97.9	580	9	US-10-090-035-7	Sequence 7, Appli
7	295	57.8	591	9	US-10-090-035-15	Sequence 15, Appli
8	277.5	54.4	436	9	US-10-090-035-21	Sequence 21, Appli
9	258	50.6	584	9	US-10-090-035-19	Sequence 19, Appli
10	258	50.6	584	9	US-10-090-035-13	Sequence 23, Appli
11	251	49.2	348	9	US-10-090-035-23	Sequence 13, Appli
12	213	41.8	237	10	US-09-923-876-2788	Sequence 2788, Ap
13	92.5	18.1	328	10	US-09-924-035A-17	Sequence 17, Appli
14	92.5	18.1	591	9	US-09-938-842A-1858	Sequence 1858, Ap
15	88.5	17.4	546	10	US-09-728-445-817	Sequence 817, App
16	83	16.3	3870	9	US-09-712-363-113	Sequence 113, App
17	82.5	16.2	3966	9	US-10-198-846-13254	Sequence 13254, A
18	80	15.7	10062	10	US-09-888-615-5	Sequence 5, Appli
19	78.5	15.4	1602	10	US-09-833-790-417	Sequence 417, App
20	78	15.3	1059	9	US-10-232-563-4	Sequence 4, Appli
21	78	15.3	1189	10	US-09-900-237-5	Sequence 5, Appli
22	78	15.3	2681	9	US-10-232-563-1	Sequence 1, Appli
23	77	15.1	1080	9	US-10-232-563-5	Sequence 5, Appli
24	77	15.1	6381	10	US-09-969-347-216	Sequence 216, App
25	76	14.9	1203	9	US-09-879-312-1	Sequence 1, Appli
26	76	14.9	2627	9	US-10-037-270-434	Sequence 434, App
27	75.5	14.8	491	9	US-09-854-133-343	Sequence 343, App
28	75.5	14.8	491	10	US-09-738-973-343	Sequence 3, Appli
29	75.5	14.8	25309	9	US-10-291-737-3	Sequence 1, Appli
30	75	14.7	8979	9	US-09-738-626-2739	Sequence 2739, Ap
31	75	14.7	3309400	9	US-10-103-313-83	Sequence 83, Appli
32	74.5	14.6	1458	9	US-10-738-626-1	Sequence 95, Appli
33	74.5	14.6	1519	9	US-10-153-668-95	Sequence 30, Appli
34	74	14.5	464	10	US-09-833-790-30	Sequence 1, Appli
35	74	14.5	558	9	US-09-998-660-1	Sequence 6743, Ap
36	74	14.5	593	10	US-09-864-761-6743	Sequence 325, App
37	74	14.5	910	12	US-10-062-254-325	Sequence 9679, Ap
38	74	14.5	32082	9	US-09-764-891-9679	Sequence 1494, Ap
39	73.5	14.4	279	9	US-09-984-130-143	Sequence 6, Appli
40	73.5	14.4	858	9	US-09-938-842A-1494	Sequence 7, Appli
41	73.5	14.4	930	12	US-10-052-798-7	Sequence 15, App
42	73.5	14.4	939	12	US-10-052-798-7	Sequence 457, App
43	73.5	14.4	1084	9	US-09-938-842A-457	Sequence 6480, Ap
44	73.5	14.4	2331	9	US-09-918-995-6480	
45	73	14.3	472	9		

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Alignment Scores:
Pred No.: 1.3e-58
Score: 510.00
Percent Similarity: 100.00%
Length: 574
Matches: 93
Conservative: 0

```
Best-Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-3 (1-574)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 96 ATGGCTTACTACGAGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGCCCGGCC 155
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 156 GGCTTCGGCCGACGCGCGCGGCTCCAGCAGCAGCTCGTCAAGAGAGTTCGAGGAG 215
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisGlyGly 60
DB 216 GTCGACAGGTCTCAGCGCGCGGCCAACCAACCACCATGGTTCACCAACGCGGC 275
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 276 CACGGCTTCGTGTCGCGAGACCAAGGTCGAGAGGACATCAACACCTGCACCGCGGC 335
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 336 GTCCACAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 374

RESULT 2
US-10-090-035-5
; Sequence 5, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(377)
US-10-090-035-5
Alignment Scores:
Pred. No.: 1,31e-58 Length: 577
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-5 (1-577)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 99 ATGGCTTACTACGAGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGCCCGGCC 159
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 159 GGCTTCGGCCGACGCGCGCGGCTCCAGCAGCAGCTCGTCAAGAGAGTTCGAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisGlyGly 60
DB 219 GTCGACAGGTCTCAGCGCGCGGCCAACCAACCACCATGGTTCACCAACGCGGC 278
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 329 GTCCACAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 367

RESULT 4
US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
US-10-090-035-4 (1-93) x US-10-090-035-1 (1-676)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 89 ATGGCTTACTACGAGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGCCCGGCC 148
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 149 GGCTTCGGCCGACGCGCGCGGCTCCAGCAGCAGCTCGTCAAGAGAGTTCGAGGAG 208
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisGlyGly 60
DB 209 GTCGACAGCGTATCAGCGCGCGGCCAACCAACCACCATGGTTCACCAACGCGGC 268
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 269 CACGGCTTCGTGTCGCGAGACCAAGGTCGAGGAGGACATCAACACCTGCACCGCGGC 328
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 329 GTCCACAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 367

RESULT 3
US-10-090-035-1
; Sequence 1, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(367)
US-10-090-035-1
Alignment Scores:
Pred. No.: 1,59e-58 Length: 676
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-1 (1-676)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 89 ATGGCTTACTACGAGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGCCCGGCC 148
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 149 GGCTTCGGCCGACGCGCGCGGCTCCAGCAGCAGCTCGTCAAGAGAGTTCGAGGAG 208
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisGlyGly 60
DB 209 GTCGACAGCGTATCAGCGCGCGGCCAACCAACCACCATGGTTCACCAACGCGGC 268
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 269 CACGGCTTCGTGTCGCGAGACCAAGGTCGAGGAGGACATCAACACCTGCACCGCGGC 328
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 329 GTCCACAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 367

RESULT 4
US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
```

;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9
;; LENGTH: 529
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (53)...(331)
US-10-090-035-9

Alignment Scores:
Pred. No.: 2,93e-58 Length: 529
Score: 507.00 Matches: 92
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.92% Mismatches: 0
Query Match: 99.41% Indels: 0
DB: Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-9 (1-529)

QY 1 MetalatYrTyGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 53 ATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 112
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 113 GGCTTCGGCCGCGCACGGCGGGGCTCCAGGAGGAGTCTCAAGGAGAGTTCGAGGAG 172
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
|||||
Db 173 GTCGACACGGTTCGACGGCGCGGGCTCCAGGAGGAGTCTCAAGGAGAGTTCGAGGAG 172
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGlyGlu 80
|||||
Db 233 CACGGCTTCGTGGTGGCGGACACCGGCTCGGAGGAGGAGTCAACACCTGCACCGCGGAG 292
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 293 GTCCACGAGCGCAGGAGAGCTTCTCCGCCAGGCGCTAAC 331

RESULT 5

US-10-090-035-17
;; Sequence 17, Application US/10090035
;; Patent No. US20020170089A1
;; GENERAL INFORMATION:
;; APPLICANT: Simmons, Carl R.
;; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
;; FILE REFERENCE: 35718/242990
;; CURRENT APPLICATION NUMBER: US/10/090,035
;; CURRENT FILING DATE: 2002-02-28
;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17
;; LENGTH: 524
;; TYPE: DNA
;; ORGANISM: Triticum aestivum
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (57)...(338)
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(524)
;; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-17

Alignment Scores:
Pred. No.: 2,85e-57 Length: 524
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0

Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: Gaps: 1

US-10-090-035-4 (1-93) x US-10-090-035-17 (1-524)

QY 1 MetalatYrTyGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 57 ATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 116
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 117 GGCTTCGGCCGCGCACGGCGGGCTCCAGGAGGAGTCTCAAGGAGAGTTCGAGGAG 176
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 59
|||||
Db 177 GTCGACACGGTTCACGGCGGGGCGGCAACACACCATGTCACACGCGC 236
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGly 79
|||||
Db 237 GGCCACGGCTTCGTGGTGGCGGAGCAGGCTCGAGGAGGAGTCAACACCTGCACCGC 296
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 297 GAGGTCACGAGCGCAGGAGAGCTTCTCCGCCAGGCGCTAAC 338

RESULT 6

US-10-090-035-7
;; Sequence 7, Application US/10090035
;; Patent No. US20020170089A1
;; GENERAL INFORMATION:
;; APPLICANT: Simmons, Carl R.
;; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
;; FILE REFERENCE: 35718/242990
;; CURRENT APPLICATION NUMBER: US/10/090,035
;; CURRENT FILING DATE: 2002-02-28
;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7
;; LENGTH: 580
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (99)...(380)
US-10-090-035-7

Alignment Scores:
Pred. No.: 3,24e-57 Length: 580
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: Gaps: 1

US-10-090-035-4 (1-93) x US-10-090-035-7 (1-580)

QY 1 MetalatYrTyGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 99 ATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 158
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 159 GGCTTCGGCCGCGCACGGCGGGCTCCAGGAGGAGTCTCAAGGAGAGTTCGAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 59
|||||
Db 219 GTCGACACGGTTCACGGCGGGCGGCAACACCATGTCACACGCGC 278
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGly 79

Db 279 GCCACGGCTTGGTGGCGAGACAGGGTCGAGGAGGACATCAACACCTGCACCGGC 338
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 339 GAGTCCAGCAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 380
RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.: 4e-30 Length: 591
Score: 295.00 Matches: 57
Percent Similarity: 75.53% Conservatives: 14
Best Local Similarity: 60.64% Mismatches: 19
Query Match: 57.84% Indels: 4
Gaps: 2
US-10-090-035-4 (1-93) x US-10-090-035-15 (1-591)
QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db 61 ATGGCTCATTACCCAGGAGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTGACCCGCCACC 120
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db 121 GCGGCTTCTCGCGCGCGCGGTGCGAGCAGCAGCATCTGTCAGGAGAGCGTTCCAG 180
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisGly 59
Db 181 GAGATCGAC-----AGTCCGGTCCGCGCGCGCCACCACCAACCAACCGGC 231
QY 60 GlyHisGlyPheValArgGluThrArgValGluGluAspLeuThrCysThrGly 79
Db 232 AACGACTACCTGATGTGCGGAGCAGCAAGGTGAGGAGGACTTCAACACCTGCACCGGC 291
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 292 GAGTCCCGCAGCGCAGCAGAGCTTCTGCTCAAGTCCGAC 333
RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990

; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(326)
US-10-090-035-21
Alignment Scores:
Pred. No.: 5.68e-28 Length: 436
Score: 277.50 Matches: 60
Percent Similarity: 73.47% Conservatives: 12
Best Local Similarity: 61.22% Mismatches: 13
Query Match: 54.41% Indels: 13
Gaps: 6
US-10-090-035-4 (1-93) x US-10-090-035-21 (1-436)
QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 54 ATGGCGCACTTCCAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGGCGGTGGCAACCCG 113
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPhe--- 38
Db 114 GCC-----CGCCGCGCGCGCGGTGCGAGGAGCAGCATCTGTCAGGAGACGTTCTGTG 164
QY 39 GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHis 58
Db 165 CAGGAGTTCGACACCTCCGCGCGC-----CGCCACGGTCCACCGGTCCACCAC 212
QY 59 Gly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspLeu 75
Db 213 GCGCGCGCGCTCTGGTGCAC---TTCGAGGTGCGCGAGAGCAGCGGTCTCGAGGACTTCAAC 269
QY 76 ThrCysThrGlyGluValHisGlyGlyValGlnGlnHisGlyArgGluSerPheLeuAlaArgAlaAsn 93
Db 270 ACCCGCAGCGGGGAGTTCACGAGCGCAAGGAGAACTTCGTCGTCAGGCGCGAT 323
RESULT 9
US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:

Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-4 (1-93) x US-10-090-035-19 (1-584)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 ATGGCGCACTTCCAGGAGTGGACTACTGCTCGAGGAGGTGAGGCGGTGGCTACCCG 105
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 GCC-----CGCCGCGCTGCGGCGCGTGCAGGAGCACATCGTCAAGGAGACGTTTC 156
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHis 57
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 GTGAGGAGTTCGACACCGCGCGCGC-----GCCAYGGTCCACCGGTAC 204
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 205 CAGCGCGCGTCTCGGTCTAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGGACATC 261
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 AACACCGCGCAGGAGTTCACGACGCAAGGAAATTCCTCTCCAGGCGCAT 318
```

RESULT 10

US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-23

Alignment Scores:
Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-4 (1-93) x US-10-090-035-23 (1-584)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 ATGGCGCACTTCCAGGAGTGGACTACTGCTCGAGGAGGTGAGGCGGTGGCTACCCG 105
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 GCC-----CGCCGCGCTGCGGCGCGTGCAGGAGCACATCGTCAAGGAGACGTTTC 156
```

```
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHis 57
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 GTGAGGAGTTCGACACCGCGCGCGC-----GCCAYGGTCCACCGGTAC 204
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 205 CAGCGCGCGTCTCGGTCTAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGGACATC 261
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 AACACCGCGCAGGAGTTCACGACGCAAGGAAATTCCTCTCCAGGCGCAT 318
```

RESULT 11

US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc_feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-13

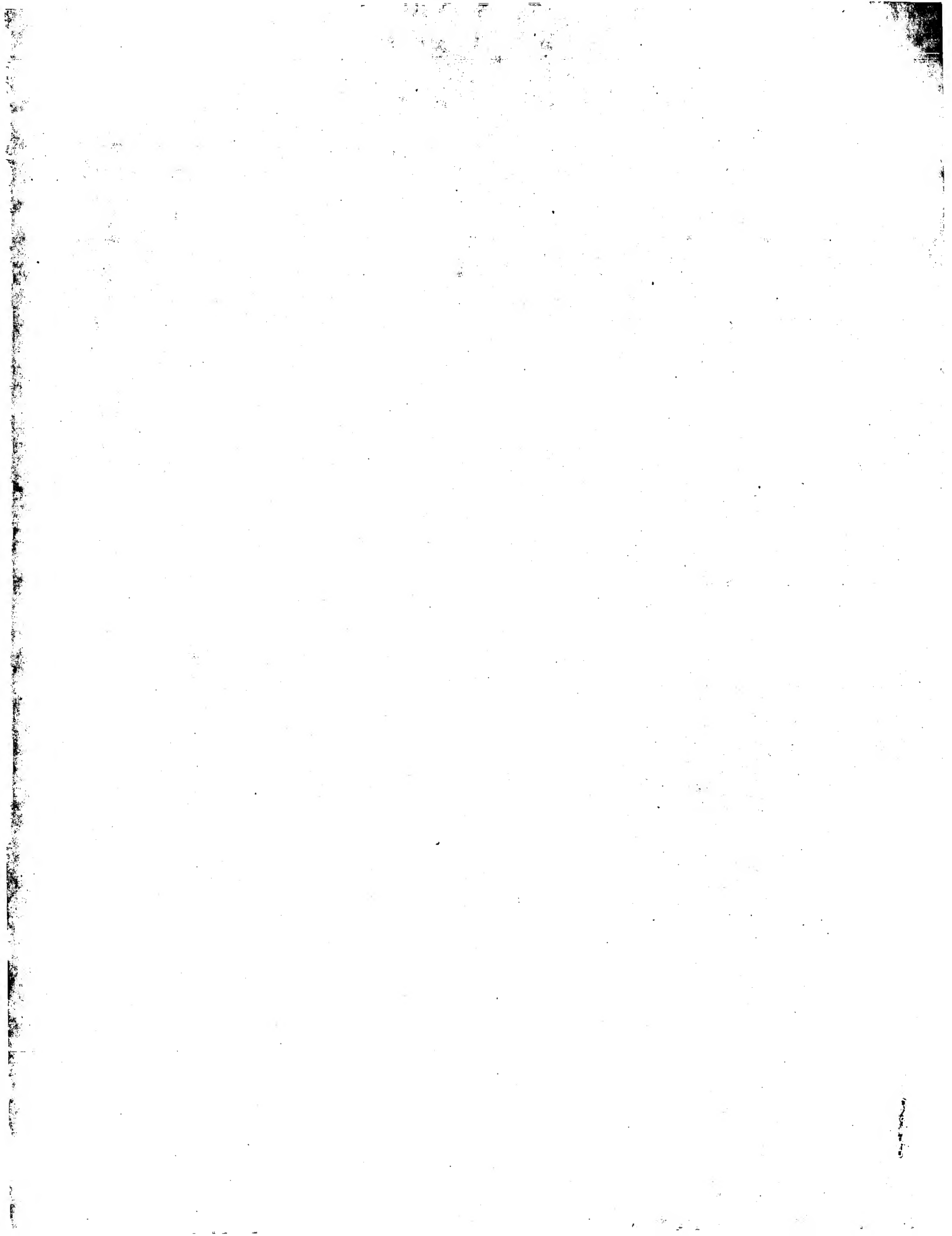
Alignment Scores:
Pred. No.: 1.39e-24 Length: 348
Score: 251.00 Matches: 51
Percent Similarity: 70.59% Conservative: 9
Best Local Similarity: 60.00% Mismatches: 21
Query Match: 49.22% Indels: 4
DB: 9 Gaps: 2

US-10-090-035-4 (1-93) x US-10-090-035-13 (1-348)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ATGGCTCACTTACCAGGAGTGGACTACTGCTCGAGGAGGTGAGGTGAGTCCGCCACCC 111
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 GCGCGTCTCTCGCGCGCGCGCGCTGCAGCACACGACGCTCGTCAAGGAGAGCTTCCAG 171
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisGly 59
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 GAGATCGACANG-----TCCGGCTCCGCGCGGCAGCAGCAGCAGCAGCAGCAGCAGC 222
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 ACCGACTACTTAAGTTCGCGAGACCAAGGTGAGGAGACTTTAACACCTCGCCGCGC 282
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 80 GluValHisGluLys 84
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 GAGTTTCGCGAGCGC 297
```

RESULT 12

US-09-923-876-2788
; Sequence 2788, Application US/09923876



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:52:33 ; Search time 29 Seconds
(without alignments)
331.081 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQHVYKFEEDVTVSAGANHHHHHHGG 93

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	93	9	US-10-090-035-2
2	510	100.0	93	9	US-10-090-035-4
3	510	100.0	93	9	US-10-090-035-6
4	507	99.4	93	9	US-10-090-035-10
5	499.5	97.9	94	9	US-10-090-035-8
6	499.5	97.9	94	9	US-10-090-035-18
7	295	57.8	91	9	US-10-090-035-16
8	277.5	54.4	91	9	US-10-090-035-22
9	273	53.5	92	9	US-10-090-035-20
10	258	50.6	92	9	US-10-090-035-24
11	245	48.0	99	9	US-10-090-035-14
12	80	15.7	359	9	US-10-232-563-7
13	80	15.7	359	10	US-09-888-615-64
14	78	15.3	352	9	US-10-232-563-2
15	77	15.1	359	9	US-10-232-563-6
16	76	14.9	400	9	US-09-879-312-2
17	73.5	14.4	285	9	US-10-278-173-16
18	73.5	14.4	309	12	US-10-052-798-9
19	73.5	14.4	312	12	US-10-052-798-10

20	73	14.3	19	9	US-09-876-904A-627	Sequence 627, App
21	73	14.3	633	10	US-09-824-735-3	Sequence 3, Appl1
22	73	14.3	633	10	US-09-801-368-338	Sequence 338, App
23	71.5	14.0	310	12	US-10-052-798-11	Sequence 11, Appl
24	71.5	14.0	354	10	US-09-925-302-501	Sequence 501, App
25	71.5	14.0	532	10	US-09-833-790-428	Sequence 428, App
26	69.5	13.6	315	9	US-10-232-563-11	Sequence 11, Appl
27	69.5	13.6	623	9	US-10-108-605-125	Sequence 125, App
28	69.5	13.6	623	9	US-10-108-605-129	Sequence 129, App
29	69.5	13.6	725	9	US-09-932-257A-19	Sequence 19, Appl
30	69	13.5	257	9	US-09-932-257A-19	Sequence 6135, Ap
31	68.5	13.4	20	9	US-09-051-013-3	Sequence 3, Appl1
32	68	13.3	139	10	US-09-813-820-8	Sequence 8, Appl1
33	67.5	13.2	118	10	US-09-205-658-120	Sequence 120, App
34	67.5	13.2	530	9	US-10-044-692-317	Sequence 317, App
35	67.5	13.2	530	9	US-10-044-539-317	Sequence 317, App
36	67.5	13.2	567	9	US-10-270-333-126	Sequence 126, App
37	67.5	13.2	605	9	US-09-741-233A-2	Sequence 2, Appl1
38	67.5	13.2	1207	9	US-10-108-605-71	Sequence 71, Appl
39	67	13.1	446	10	US-09-853-386-69	Sequence 69, Appl
40	67	13.1	507	9	US-09-795-927-10	Sequence 10, Appl
41	67	13.1	507	10	US-09-738-897-2	Sequence 201, App
42	67	13.1	608	9	US-09-975-719-201	Sequence 201, App
43	67	13.1	639	9	US-09-975-719-200	Sequence 199, App
44	67	13.1	643	9	US-09-975-719-199	Sequence 70, Appl
45	67	13.1	643	10	US-09-853-386-70	

ALIGNMENTS

RESULT 1
US-10-090-035-2
; Sequence 2, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

Query Match Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQHVYKFEEDVTVSAGANHHHHHHGG 60
DB 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQHVYKFEEDVTVSAGANHHHHHHGG 60
QY 61 HGFVVRTRVEEDINTCTGVEHRRSFARAN 93
DB 61 HGFVVRTRVEEDINTCTGVEHRRSFARAN 93

RESULT 2
US-10-090-035-4
; Sequence 4, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof

```
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-4

Query Match          100.0%; Score 510; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 3
US-10-090-035-6
; Sequence 6, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-6

Query Match          100.0%; Score 510; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 4
US-10-090-035-10
; Sequence 10, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
US-10-090-035-10

Query Match          99.4%; Score 507; DB 9; Length 93;
Best Local Similarity 98.9%; Pred. No. 1e-44;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 5
US-10-090-035-8
; Sequence 8, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-8

Query Match          97.9%; Score 499.5; DB 9; Length 94;
Best Local Similarity 98.9%; Pred. No. 5.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 59
Db 1 MAYQEVYDVCSEEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 60 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 6
US-10-090-035-18
; Sequence 18, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
US-10-090-035-18
```

```
; ORGANISM: Triticum aestivum
US-10-090-035-18

Query Match          97.9%; Score 499.5; DB 9; Length 94;
Best Local Similarity 98.9%; Pred. No. 5.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAYYQEVYDVCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MAYYQEVYDVCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 60

QY 60 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 7
US-10-090-035-16
; Sequence 16, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-090-035-16

Query Match          57.8%; Score 295; DB 9; Length 91;
Best Local Similarity 60.6%; Pred. No. 4e-23;
Matches 57; Conservative 14; Mismatches 19; Indels 4; Gaps 2;

QY 1 MAYYQEVYDVCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MAYYQEVYDVCSEEVRSVPTGFLGRGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 57

QY 60 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 58 NDYLMVRTRVEEDINTCTGEFRERKQSLKSD 91

RESULT 8
US-10-090-035-22
; Sequence 22, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-22

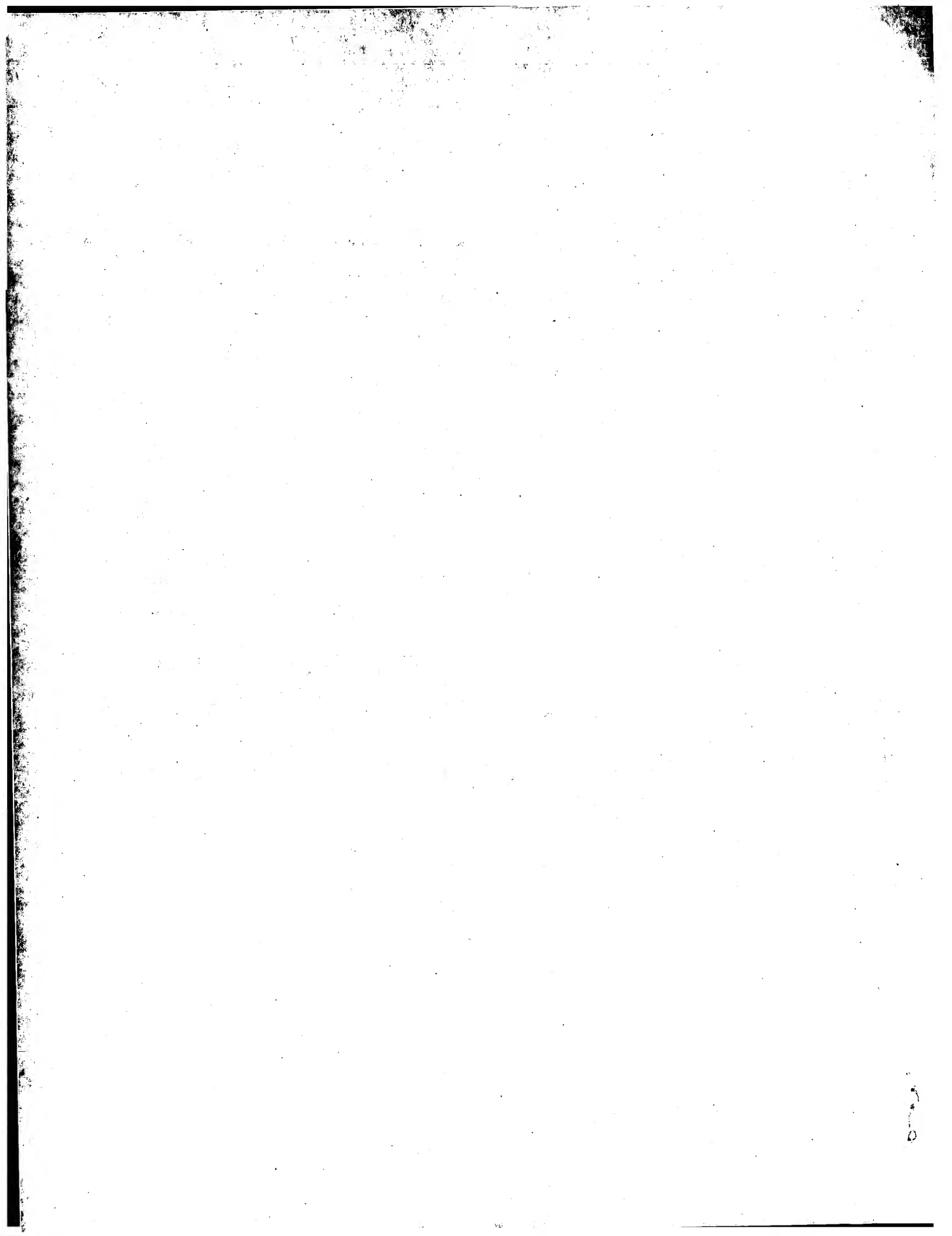
Query Match          54.4%; Score 277.5; DB 9; Length 91;
Best Local Similarity 61.2%; Pred. No. 2.4e-21;
Matches 58; Conservative 14; Mismatches 13; Indels 14; Gaps 7;

QY 1 MAYYQEVYDVCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MAHFQEVYDVCSEEVRAVGYPAGGCGGVQEHIVKETFVQEFDTAGR----RHGHHG 53

Query Match          50.6%; Score 258; DB 9; Length 92;
Best Local Similarity 58.6%; Pred. No. 2.4e-19;
Matches 58; Conservative 14; Mismatches 13; Indels 14; Gaps 7;

QY 1 MAYYQEVYDVCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDTVSRAGANHHHHHHG 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MAHFQEVYDVCSEEVRAVGYPAGGCGGVQEHIVKETFVQEFDTAGR----RHGHHG 53
```


Search completed: June 6, 2003, 13:02:10
Job time : 29 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:49:03 ; Search time 18 Seconds
(without alignments)
152.018 Million cell updates/sec

Title: US-10-090-035-4

Perfect score: 510

Sequence: 1 MAYQEDVDCSEEVSRVAP.....INTCTGVEHRESFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap:*
- 5: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pap:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	15.5	414	5	PCT-US92-06840-2
2	77	15.1	763	2	US-08-677-862-2
3	77	15.1	763	2	US-09-252-571-2
4	77	15.1	763	3	US-09-434-065-2
5	77	15.1	763	4	US-08-789-275-4
6	77	15.1	763	4	US-08-789-275-5
7	76	14.9	400	4	US-09-086-010-2
8	73.5	14.4	40	2	US-08-273-146-39
9	73.5	14.4	309	4	US-09-079-029-9
10	73.5	14.4	312	4	US-09-079-029-10
11	73	14.3	617	1	US-08-137-614A-26
12	73	14.3	633	4	US-08-557-006C-43
13	73	14.3	637	3	US-08-072-064-1
14	73	14.3	637	3	US-08-072-064-4
15	73	14.3	637	3	US-08-072-064-6
16	73	14.3	637	3	US-08-072-064-8
17	73	14.3	637	5	PCT-US92-08558-1
18	72.5	14.2	79	2	US-08-448-418-97
19	71.5	14.0	310	4	US-09-079-029-11
20	70.5	13.8	284	3	US-09-184-658-40
21	70	13.7	398	4	US-09-461-474-17
22	69.5	13.6	726	4	US-09-126-980-2
23	69.5	13.6	726	4	US-09-476-482-2
24	68.5	13.4	281	4	US-09-517-605-6
25	68.5	13.4	281	4	US-09-423-439-44
26	68	13.3	139	4	US-08-856-253-8
27	67.5	13.2	379	4	US-09-457-040B-36

28	67.5	13.2	530	4	US-08-974-549A-603	Sequence 603, App
29	67	13.1	608	4	US-09-199-637A-201	Sequence 201, App
30	67	13.1	639	4	US-09-199-637A-200	Sequence 200, App
31	67	13.1	643	4	US-09-199-637A-199	Sequence 199, App
32	67	13.1	645	4	US-09-199-637A-253	Sequence 253, App
33	66.5	13.0	449	2	US-08-927-394-2	Sequence 2, Appli
34	66	12.9	159	3	US-08-991-890-4	Sequence 4, Appli
35	66	12.9	349	4	US-09-011-769A-47	Sequence 47, Appli
36	66	12.9	349	4	US-09-011-769A-60	Sequence 60, Appli
37	66	12.9	349	4	US-09-011-769A-64	Sequence 64, Appli
38	66	12.9	379	1	US-08-279-270A-1	Sequence 1, Appli
39	65.5	12.8	474	2	US-08-650-000-4	Sequence 4, Appli
40	65.5	12.8	474	4	US-09-042-785A-8	Sequence 8, Appli
41	65.5	12.8	474	6	5395760-4	Patent No. 5395760
42	65.5	12.8	705	2	US-08-770-761A-7	Sequence 7, Appli
43	64.5	12.6	2353	4	US-08-984-709A-50	Sequence 50, Appli
44	64	12.5	60	1	US-08-255-457-1	Sequence 1, Appli
45	64	12.5	60	2	US-09-115-032-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
PCT-US92-06840-2
; Sequence 2, Application PC/TUS9206840
; GENERAL INFORMATION:
; APPLICANT: Shi, Yang
; APPLICANT: Seto, Edward
; APPLICANT: Shenk, Thomas
; TITLE OF INVENTION: Y1 TRANSCRIPTION FACTOR AND METHODS OF
; TITLE OF INVENTION: ISOLATING SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas - 7th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06840
; FILING DATE: 19920814
; CLASSIFICATION:
; PRIOR APPLICATION: AU 1805
; APPLICATION DATA: US 07/746,485
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Dennis, Manette
; REGISTRATION NUMBER: 30,623
; REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-06840-2

Query Match 15.5% Score 79; DB 5; Length 414;
Best Local Similarity 29.2%; Pred. No. 0.088;
Matches 21; Conservative 6; Mismatches 25; Indels 20; Gaps 3;

```
QY 21 GFGRRGGVQQRHVKEEVEDTVSRAGANHHHHHHGGH-GFVVRTRVEEDINTCTG 79
Db 55 GGDHGGG-----GGHGHAGHHHHHHHHPPMIALQPLVTD-----PT. 95
QY 80 EVHERRESFLAR 91
Db 96 QVHHQEVILQ 107

RESULT 2
US-08-677-862-2
; Sequence 2, Application US/08677862
; Patent No. 5874230
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,862
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-252-571-2
; Query Match 15.1%; Score 77; DB 2; Length 763;
; Best Local Similarity 33.3%; Pred. No. 0.33;
; Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
QY 11 SEEVRSVAPAGFRHGG-----GVOQHVVKEEVEDTVSRA-----GANHHHHHG 56
Db 555 SPQVRQQFPAPLWGSGETAPQVTVEHPVQETTFHVAPOQNALHHHHGSSHHHHHHH 614
QY 57 HHGGHG 62
Db 615 HHHHHG 620

RESULT 4
US-09-434-065-2
; Sequence 2, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/434,065
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```


RESULT 6
US-08-789-275-5
; Sequence 5, Application US/08789275A
; Patent No. 6231664
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```


APPLICATION NUMBER: US/09/079,029
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1101R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5416
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-079-029-10

Query Match 14.4%; Score 73.5; DB 4; Length 312;
Best Local Similarity 32.9%; Pred. No. 0.29;
Matches 25; Conservative 6; Mismatches 14; Indels 31; Gaps 6;
QY 6 EVD-YGSEVRVSYPAGFGRH---GGVVOHVVKKEFEVDVSRAGANHHHHHHGHHG 61
Db 259 EADYCNRSDDSS-----GNHVFGGTYKL-----TVLGAAHHHHHHHHG----- 296
QY 62 GFVRETRV--EEDIN 75
Db 297 ---AAEQKLISEDLN 309

RESULT 11
US-08-137-614A-26
Sequence 26, Application US/08137614A
Patent No. 5487976

GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Knipple, Douglas C.
APPLICANT: Henderson, Joseph E.
TITLE OF INVENTION: Gene Encoding An Insect
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,614A
FILING DATE: 15-OCT-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-137-614A-26

Query Match 14.3%; Score 73; DB 1; Length 617;

Best Local Similarity 27.8%; Pred. No. 0.78;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;
QY 17 VAPAGFGRHGGVQHVVKKEFEVDVSRAGANHHHHHHHHG---GHGFVYRET----- 68
Db 413 VPGGPGGPGGGV-----NVGVGMGMGPEHGHGHHHSHGHAPKQTVSNRP 462
QY 69 ----RVEEDINT 76
Db 463 IGFSNIQNVGT 474

RESULT 12
US-08-557-006C-43
Sequence 43, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Beri, Rajinder K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PHM37588/UST
CURRENT APPLICATION NUMBER: US/08/557,006C
CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 43
LENGTH: 633
TYPE: PRT
ORGANISM: Yeast
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(633)
OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match 14.3%; Score 73; DB 4; Length 633;
Best Local Similarity 76.9%; Pred. No. 0.81;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 50 NHHHHHHHHGHHG 62
Db 21 HHHHHHHHHHHG 33

RESULT 13
US-08-072-064-1
Sequence 1, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; ORGANISM: Drosophila melanogaster
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: III; polytene subregion 66F
; MAP POSITION: approximately map unit 26
; US-08-072-064-1

Query Match 14.3%; Score 73; DB 3; Length 637;
Best Local Similarity 27.8%; Pred. No. 0.81;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;

QY 17 VAPAGFRHGGVQVHVVKKEFEVDTVSRAGANHHHHHHG---GHGFVVRET----- 68
DB 434 VPGGPGGPGGGV-----NVGVGMGMPHGHHGHAHSHGHPAKQTVSNRP 483
QY 69 ----RVEEDINT 76
DB 484 IGFSNIQNVGT 495

RESULT 15
US-08-072-064-6
; Sequence 6, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-072-064-6

Query Match 14.3%; Score 73; DB 3; Length 637;
Best Local Similarity 27.8%; Pred. No. 0.81;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;

QY 17 VAPAGFRHGGVQVHVVKKEFEVDTVSRAGANHHHHHHG---GHGFVVRET----- 68
DB 434 VPGGPGGPGGGV-----NVGVGMGMPHGHHGHAHSHGHPAKQTVSNRP 483
QY 69 ----RVEEDINT 76
DB 484 IGFSNIQNVGT 495

RESULT 14
US-08-072-064-4
; Sequence 4, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
```

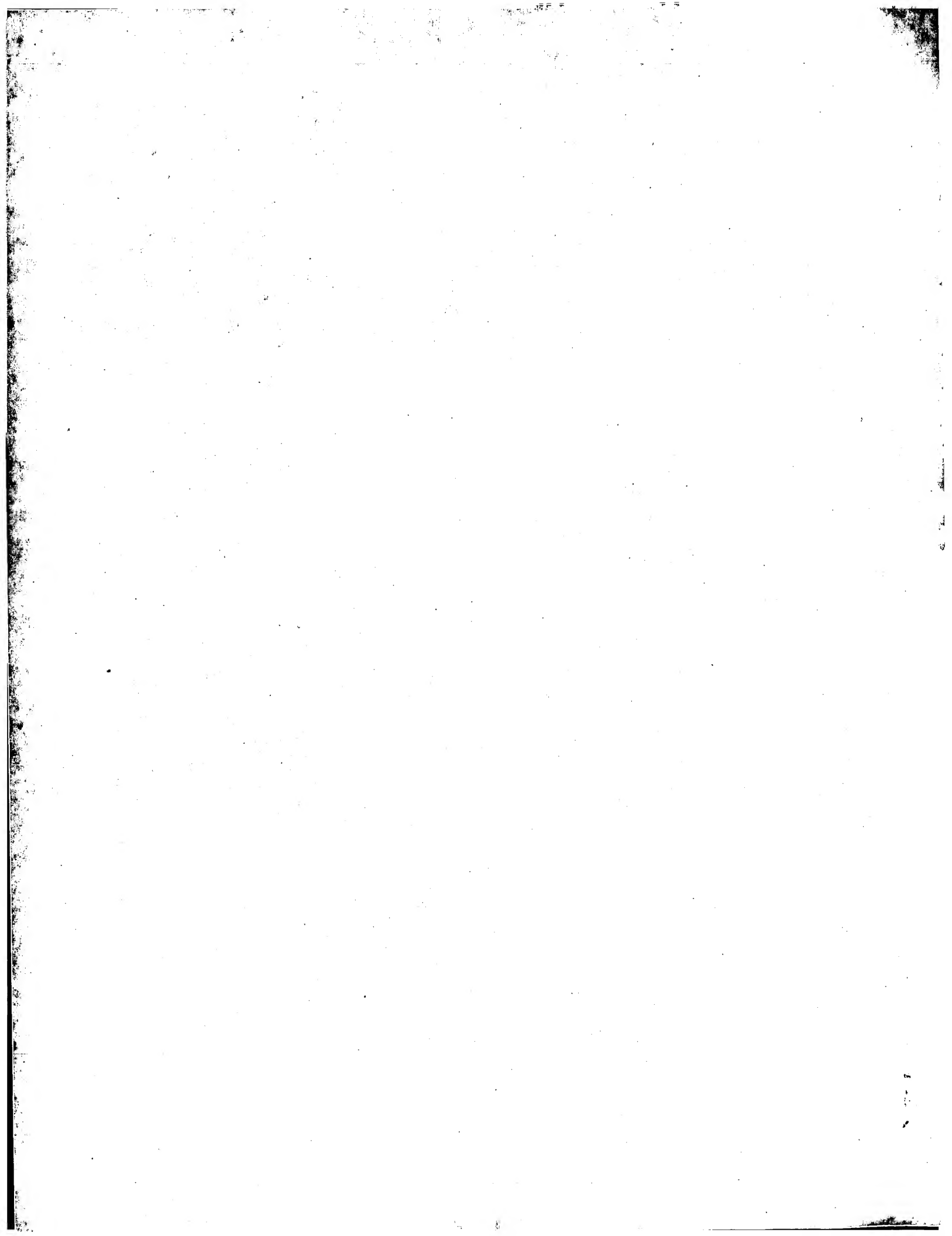
Mon Jun 16 14:55:57 2003

us-10-090-035-4.ra1

Page 7

Oy 69 ----RVEDINT 76
Db 484 IGFSNIOOVGT 495
::: |

Search completed: June 6, 2003, 12:54:08
Job time : 19 secs



Result No.	Query No.	%		Length	DB	ID	Description
		Score	Match				
1	84	16.5	441	21	AGC26460	Arabidopsis thalia	
2	84	16.5	473	21	AGC26459	Arabidopsis thalia	
3	84	16.5	488	21	AGC26458	Arabidopsis thalia	
4	84	16.5	1561	22	AB866445	Drosophila melanog	
5	83	16.3	168	22	AB863257	Drosophila melanog	
6	81.5	16.0	271	23	ABG66039	scfv lb4 antibody	
7	80.5	15.8	347	22	AB867674	Drosophila melanog	
8	80.5	15.8	384	21	AGC13903	Arabidopsis thalia	
9	80.5	15.8	388	21	AGC13902	Arabidopsis thalia	
10	80.5	15.8	388	21	AGC49725	Arabidopsis thalia	

ALIGNMENTS

16-APR-1999; 99US-0129845.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

us-10-090-035-4.rag

Mon Jun 16 14:55:57 2003

PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134219.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134221.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	05-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	06-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.

PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	27-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151085.
PR	27-AUG-1999;	99US-0151086.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	25-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 16.5%; Score 84; DB 21; Length 473;		
Best Local Similarity 36.8%; Pred. No. 0.07; Indels 8; Gaps 3;		
Matches 21; Conservative 10; Mismatches 18;		
QY	8	DYCEEVRSVAPAGFGRHGGVQOHVYKKEFEVDVTSRAGANHHHHHGH---HGGH 61
DB	409	DYIEFV-SLLPKSIRR----VAEPIPEEVQKVLLEAKAGDDHDDHHHGHHAHAGY 460
RESULT 3		
AAG26458		
ID	AAG26458	standard; Protein; 488 AA.
XX		
AC	AAG26458;	
XX		
DT	17-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 30921.
XX		
KW	Protein identification;	signal transduction pathway; metabolic pathway;
KW	hybridisation assay;	genetic mapping; gene expression control; promoter;
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000;	2000EP-0301439.
XX		
XX	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135134.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 25-AUG-1999; 99US-0149920.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

[illegible]

Mon Jun 16 14:55:57 2003

XX	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
OS		PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	02-JUL-1999;	99US-0142154.
XX		PR	06-JUL-1999;	99US-0142390.
PD	06-SEP-2000.	PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
PF	25-FEB-2000; 2000EP-0301439.	PR	13-JUL-1999;	99US-0143542.
XX		PR	14-JUL-1999;	99US-0143624.
PR		PR	15-JUL-1999;	99US-0144005.
PR		PR	16-JUL-1999;	99US-0144085.
PR		PR	16-JUL-1999;	99US-0144086.
PR		PR	19-JUL-1999;	99US-0144325.
PR		PR	19-JUL-1999;	99US-0144331.
PR		PR	19-JUL-1999;	99US-0144332.
PR		PR	19-JUL-1999;	99US-0144333.
PR		PR	19-JUL-1999;	99US-0144334.
PR		PR	19-JUL-1999;	99US-0144335.
PR		PR	20-JUL-1999;	99US-0144352.
PR		PR	20-JUL-1999;	99US-0144632.
PR		PR	20-JUL-1999;	99US-0144884.
PR		PR	21-JUL-1999;	99US-0144814.
PR		PR	21-JUL-1999;	99US-0145086.
PR		PR	21-JUL-1999;	99US-0145088.
PR		PR	22-JUL-1999;	99US-0145085.
PR		PR	22-JUL-1999;	99US-0145087.
PR		PR	22-JUL-1999;	99US-0145089.
PR		PR	23-JUL-1999;	99US-0145192.
PR		PR	23-JUL-1999;	99US-0145145.
PR		PR	23-JUL-1999;	99US-0145218.
PR		PR	23-JUL-1999;	99US-0145224.
PR		PR	26-JUL-1999;	99US-0145276.
PR		PR	27-JUL-1999;	99US-0145913.
PR		PR	27-JUL-1999;	99US-0145918.
PR		PR	28-JUL-1999;	99US-0145919.
PR		PR	02-AUG-1999;	99US-0146386.
PR		PR	02-AUG-1999;	99US-0146388.
PR		PR	02-AUG-1999;	99US-0146389.
PR		PR	03-AUG-1999;	99US-0147038.
PR		PR	04-AUG-1999;	99US-0147204.
PR		PR	04-AUG-1999;	99US-0147302.
PR		PR	05-AUG-1999;	99US-0147192.
PR		PR	05-AUG-1999;	99US-0147260.
PR		PR	06-AUG-1999;	99US-0147303.
PR		PR	06-AUG-1999;	99US-0147416.
PR		PR	09-AUG-1999;	99US-0147493.
PR		PR	09-AUG-1999;	99US-0147935.
PR		PR	10-AUG-1999;	99US-0148171.
PR		PR	11-AUG-1999;	99US-0148319.
PR		PR	12-AUG-1999;	99US-0148341.
PR		PR	13-AUG-1999;	99US-0148565.
PR		PR	13-AUG-1999;	99US-0148684.
PR		PR	16-AUG-1999;	99US-0149368.
PR		PR	17-AUG-1999;	99US-0149175.
PR		PR	18-AUG-1999;	99US-0149426.
PR		PR	20-AUG-1999;	99US-0149722.
PR		PR	20-AUG-1999;	99US-0149723.
PR		PR	20-AUG-1999;	99US-0149929.
PR		PR	23-AUG-1999;	99US-0149902.
PR		PR	23-AUG-1999;	99US-0149930.
PR		PR	25-AUG-1999;	99US-0150366.
PR		PR	26-AUG-1999;	99US-0150884.
PR		PR	27-AUG-1999;	99US-0151065.
PR		PR	27-AUG-1999;	99US-0151066.
PR		PR	27-AUG-1999;	99US-0151080.
PR		PR	30-AUG-1999;	99US-0151303.
PR		PR	31-AUG-1999;	99US-0151438.
PR		PR	01-SEP-1999;	99US-0151930.
PR		PR	07-SEP-1999;	99US-0152363.
PR		PR	10-SEP-1999;	99US-0153070.

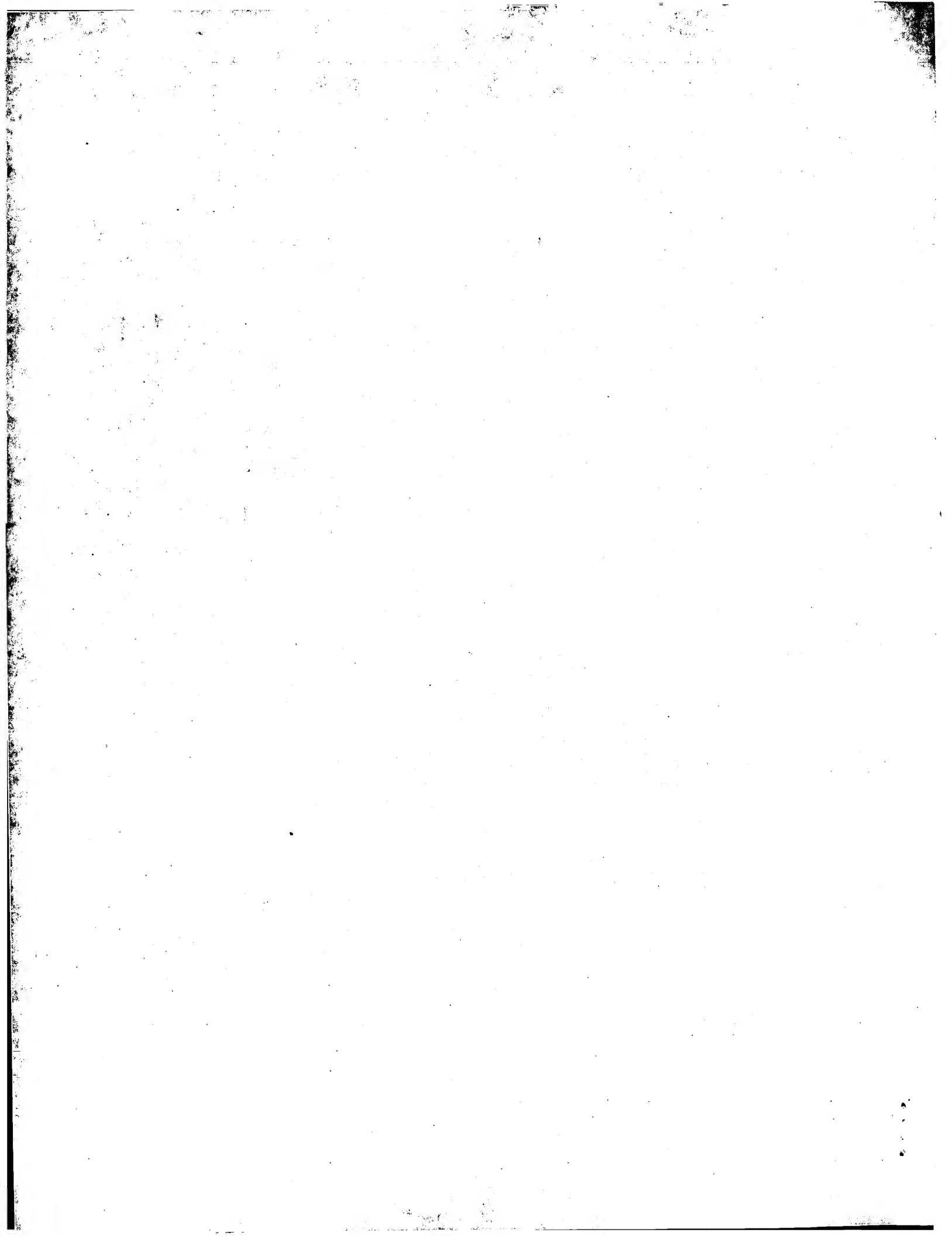
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135251.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

—

[illegible]

PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	16-AUG-1999;	9905-0148684;
PR	17-AUG-1999;	9905-0149368;
PR	18-AUG-1999;	9905-0149175;
PR	20-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	23-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	23-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151348;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	24-SEP-1999;	9905-01553486;
PR	28-SEP-1999;	9905-0155659;
PR	29-SEP-1999;	9905-0156458;
PR	04-OCT-1999;	9905-0156596;
PR	05-OCT-1999;	9905-0157117;
PR	06-OCT-1999;	9905-0157753;
PR	07-OCT-1999;	9905-0157865;
PR	08-OCT-1999;	9905-0158029;
PR	12-OCT-1999;	9905-0158232;
PR	13-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0158293;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159295;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	18-OCT-1999;	9905-0159638;
PR	21-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160980;
PR	22-OCT-1999;	9905-0160981;
PR	25-OCT-1999;	9905-0161404;
PR	25-OCT-1999;	9905-0161405;

Search completed: June 6, 2003, 12:50:08
Job time : 32.5 secs




```

; OTHER INFORMATION: H37RV
US-09-103-840A-1
Query Match          9.3%; Score 53.2; DB 4; Length 4411529;
Best Local Similarity 54.6%; Pred. No. 0.028;
Matches 106; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 149 CCCGCCGGCTTCGGCCGCCACGGCGCGCGTCCAGCAGCAGCGTGTCTCAAGGAGAGTT 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932728 CCCACCGGCATCGCGGCACCGCGGCACCGCGGCAGGAAAGGGCGCGCGCGG 3932787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 CGAGGAGGTGACAGGCTCTCACGGCGCGCGGCCAACCCACCACCATGTCACCA 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932788 CGGCGAGGGCGGTAGCAGCGGTCCGCGCGCGCGCACCAACGGCTCTGGTGGCGCTGGCGGCAC 3932847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 CGGCGGCACCGCTTCGTGTCGCGAGACAGCGGTCTCAAGAGGACATCAACACCTGCAC 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932848 CGGCGGACAGCGCGCGCGCGCGCGTGGCGGGCGCGTGGCGGGCGCGCGATACCCACCGCAT 3932907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 CGGCGAGGTCCACG 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932908 CGGCGCGCGCGCG 3932921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-476-519-10
; Sequence 10, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
US-08-476-519-10
```

```

343 AGCGAGGAGAGCTTCTCGCCAGGGCT 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 ACAGCAGCAGCAGCAGCGCATGGTT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          9.3%; Score 53.2; DB 4; Length 4403765;
Best Local Similarity 53.3%; Pred. No. 0.028;
Matches 112; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 147 GCCCGCGCGCTTCGGCCGCCACGGCGCGGTCCAGCAGCAGCTGTCAAGGAGAG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191569 GGCACCGCGCGCTTCGGCGCGGTCCAGCAGCAGCTGTCAAGGAGAG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 TTCAGGAGGTGACAGGCTCTACGGCGCGCGGCCAACCCACCACCATGTCAC 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191509 GGCAGGCGCGCGACACCGCGCGCGGCCGAGCGCGCGCGCGCGCGTGGC 1191450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 CACGCGCGCCACGCTTCGTGTCGCGAGACAGGCTCGAAGAGGACATCAACACCTGC 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191449 CACGCGCGGTTCGGCGCGGTGCTGTCGCGCAACGCGCGCGCGCGCGCGCGC 1191390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 ACCGCGGAGTTCACAGCGCGGAGGAGC 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191389 GCGCGCAACGCGCGCAACGCGCGCACGGC 1191360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
```

Query Match	9.0%;	Score 51.8;	DB 1;	Length 2244;
Best Local Similarity	54.5%;	Pred. No. 0.025;		
Matches 104;	Conservative 0;	Mismatches 87;	Indels 0;	Gaps 0;
QY	151	CGGCCGGCTTCGGCCGCACGGCGCGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTCG	210	
Db				
QY	1877	CGGCCAGCAGACCAACGGCAAGCTGATGGAGCAGTTGCGCTGGTTCAAGCCGACGCGC	1936	
Db				
QY	211	AGGAGTTCGACAGGTTCTACGCGCGCGGCCCAACCAACCACCACTCATGGTCAACCACG	270	
Db				
QY	1937	CGCAGGCGCAGCGGCTACTTCAACGGCGCCGACAAACACGCGCTGGCGTGGCGCATCG	1996	
Db				
QY	271	CGCGCCACGCGCTTCGTGGTCGGGACGACAGGCTCGAAGAGGACATCAACACCTGCACCG	330	
Db				
QY	1997	ACGGCAGCGAGTTCGGCGACAGCGCCAGCGCGATCTACGTCGCTACACGGCTGGTCCG	2056	
Db				
QY	331	CGCAGTTCAC	341	
Db				
QY	2057	CGCGGTGCAC	2067	
Db				

RESULT 5
 PCT-US95-09323-10
 ; Sequence 10, Application PC/TUS9509323
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
 ; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
 ; NUMBER OF SEQUENCES: 11
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/09323
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/281902
 ; FILING DATE: 28-JUL-1994
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2244 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..2241
 PCT-US95-09323-10

Query Match	9.0%	Score 51.8;	DB 5;	Length 2244;
Best Local Similarity	54.5%;	pred. No. 0.025;		
Matches 104;	Conservative	0;	Mismatches 87;	Indels 0;
Gaps 0;				
QY	151	CGCCCGGCTTCGGCCGCCACGGCGGGCGGTCCAGCAGCAGCTGTCGAAGGAAAGTTTCG	210	
Db	1877	CGGCCAGCGACCAACGGAACGCTGATGAGCAGTTGCGCTGGTTCAAGCCCGAGGGCG	1936	
QY	211	AGGAGTCGACAGGCTCTCACGCGCGGGGGCCAAACCACCACCACTGTCACCACG	270	
Db	1937	CGCAGCCGACAGCGGCTACTTCAAGCGCGCGACAAACAGCCCTTGGCTGGCGCATCG	1996	
QY	271	GGCGCCACCGGCTTCGTGGTGGCGGACACAGGTCGAAGAGACATCAACACCTGCACCG	330	
Db	1997	ACGGCAGCAGTTTCGGCGACAGCGCCAGCGCATCTACGTCGCCCTACAAACGGCTGTCGG	2056	
QY	331	GGCAGGTCCAC	341	
Db	2057	GGCGGCTGCAC	2067	

RESULT 7
PCT-US95-09323-1
; Sequence 1, Application PC/TUS9509323
; GENERAL INFORMATION:
; APPLICANT:

```

RESULT 6
US-08-476-519-1
; Sequence 1, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-476-519-1

```

Query Match	9.0%;	Score 51.8;	DB 1;	Length 2334;
Best Local Similarity	54.5%;	pred. No. 0.025;		
Matches 104;	Conservative 0;	Mismatches 87;	Indels 0;	Gaps 0;
QY	151	CGCGCGCTTCGGCCGCCACCGCGCGCGCTCCAGCAGCAGCTGCTCAAGAGAGAGTTCCG	210	
Db	1970	CGGCCAGCAGCAACCGGCAAGCTGATGAGGAGATTGCCGCTGGTTCAAGCCCGACGCGC	2029	
QY	211	AGGAGTCGACCGGTCTTCACGCGCGCGGCCCAACCACCACCACCATGTGTCAACCACG	270	
Db	2030	CGCAGGCGCAGACGCGCTACTTCAAGGCGCCGACAAACACGCGCTGGCGGCGATCG	2089	
QY	271	GCSCGCCACGCTTCGTGGTGGCGGAGACACGAGGTTCGAGAGAGCATCAACACCTGCACCG	330	
Db	2090	ACGCGACGAGTTTCGGCGGACAGCCGACGCGCATCTACGTTCGCTTCAACGGGTGGTCCG	2149	
QY	331	GGGAGGTCCAC	341	
Db	2150	GGCGGTGCAC	2160	

RESULT 7
PCT-US95-09323-1
; Sequence 1, Ap
; GENERAL INFOR
; APPLICANT:

TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
 TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
 NUMBER OF SEQUENCES: 11
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/09323
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/281902
 FILING DATE: 28-JUL-1994
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2334 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PCT-US95-09323-1
 Query Match 9.0%; Score 51.8; DB 5; Length 2334;
 Best Local Similarity 54.5%; Pred No. 0.025; Indels 0; Gaps 0;
 Matches 104; Conservative 0; Mismatches 87;
 QY 151 CGCGCGGGCTTCGCGCGCCACGCGCGGCGTCCAGCAGCAGCTGCTCAAGAGAGATTGCG 210
 DB 1970 CGGCCAGCGACACCAACGCAACGTGTGAGCAGCTTGGCTTCAAGCCCGACGGCG 2029
 QY 211 AGAGGTGTCAGCGGTCTCACGCGCGCGCCCAACACACACCATGTGTCACCAAG 270
 DB 2030 CGCAGGCGCAGCAGCGGCTACTTCAAGGGCGCGCGACACCGCCCTGGCCTGGCGCATCG 2089
 QY 271 GCGGCGCACGGCTTCGTGGTGGCGAGACACGAGGTGGAAGAGACATCAACACCTGCACCG 330
 DB 2090 ACGGCAGCGAGTTCGCGCAGCGCGACAGCGCGATCTACGTGCGCTACAAACGGCTGGTCCG 2149
 QY 331 GCGAGGTCCAC 341
 DB 2150 GCGCGGTGCAC 2160
 RESULT 8
 US-09-103-840A-1/c
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37RV
 US-09-103-840A-1
 Query Match 8.9%; Score 51; DB 4; Length 4411529;
 Best Local Similarity 52.6%; Pred. No. 0.078;
 Matches 111; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 132 CACGTGAGTGGGGTCCGCGCGCTTCGCGCGCCACGCGCGGCGTCCAGCAGCAC 191

.....

Mon Jun 16 14:55:57 2003

```
NAME/KEY: polyA_site
LOCATION: 7314..7319
FEATURE:
NAME/KEY: prim_transcript
LOCATION: 2072
FEATURE:
NAME/KEY: promoter
LOCATION: 1..2155
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 2042..2049
FEATURE:
NAME/KEY: mRNA
LOCATION: 2153..7079
US-08-250-848-2

Query Match      8.6%; Score 49.2; DB 2; Length 7559;
Best Local Similarity 49.6%; Pred. No. 0.094; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 128;

QY 110 GGAGGTGACTACTGTCGAGAGAGTGTGAGTGTGCGTGGCCCGCGGCTTCGGCCGCCA 169
DB 2490 GAAGCTGGGGAGCTGGCGCCAAAGCTACAGGGGTGGCCCGCGCAGCCCATCTCGT 2549

QY 170 CGCGCGGGGTCCACAGCAGCTGCTCAAGAGAGTTCGAGGAGGTGCACACGCTTC 229
DB 2550 GCGAGAGTCCATCTGCACATGCTCAACCTCGCCCAACCTGGCGGAGGAGTGCATGCC 2609

QY 230 ACAGCGGGCGCCCAACCAACCAACCATGTGTACACAGCGGCGCCAGGCTTCGTGT 289
DB 2610 GCACCGCGCGCGCAACAGCAAGCTCAAGAAAGTGGTTCGCGCAGGAGGTTCGCCAC 2669

QY 290 GCGGAGACAGGTCGAGAGACATCAACACCTGCACCGGCGGAGGTCCAGCGGCGAG 349
DB 2670 CACCGAGTCCGACATCGAGAGAGCTCAAGCGCTCGTGTGCGAGGTTCGCGAAGTCC 2729

QY 350 GGAGAGTTCCTCG 363
DB 2730 CGAGAGGTGTTCC 2743
```

```
RESULT 11
US-08-440-856A-1
; Sequence 1, Application US/08440856A
; Patent No. 5750873
; GENERAL INFORMATION:
; APPLICANT: DELLAPORTA, STEPHEN L.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
; TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE. N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.856A
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A.
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 05463-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1517
; TELEFAX: (202) 887-0763
```

```
TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-440-856A-1

Query Match      8.5%; Score 48.6; DB 1; Length 1236;
Best Local Similarity 49.4%; Pred. No. 0.1; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 129;

QY 116 GGACTACTGTCGAGGAGGTGTGAGTGTGCGTGGCCCGCGGCTTCGCGCCGCGG 175
DB 743 GAACCGCGCTTCGAGCTGCGCGCACGCGGTCCGGGTCAACTGCTCTCGCCCTTCGG 802

QY 176 CGCGGTCCAGCAGCAGCTGTGTCAGAGAGAGTTCGAGGAGGTTCACACAGGTCTC 235
DB 803 CGTCCGACGCCCATGTCTCATCAACGCTGCGCGCCAGGCCACGACGCCACGCCGA 862

QY 236 CGCGGCCAACCAACCAACCAACCATGTGTACACAGCGGCGCCAGGCTTCGTGT 295
DB 863 CGCGCAGCAGACCTCGACCTCGACCTGCGCGCCAGGCCACGACGCCAGGAGTGA 922

QY 296 GACCAGGTCGAAGAGGAGATCAACCTGTACCGCGGAGGTTCACAGCGCAGGAGAG 355
DB 923 GAAGATGAGGAGGTGTGTCAGGGGCTTGCGCCACGCTCAAGGGCCCCACGCTC 982

QY 356 CTCTCTCGCCAGGC 370
DB 983 GGACATCGCCGAGGC 997
```

```
RESULT 12
US-08-133-711-44
; Sequence 44, Application US/08133711
; Patent No. 5525463
; GENERAL INFORMATION:
; APPLICANT: Zolig, Werner
; TITLE OF INVENTION: Methods and reagents for detection of
; TITLE OF INVENTION: pathogens using superoxide dismutase gene
; TITLE OF INVENTION: targeting
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: U.S.A.
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,711
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92810780.4
; FILING DATE: 13-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Patricia S. Rocha
; REGISTRATION NUMBER: 31,054
; REFERENCE/DOCKET NUMBER: 4095/95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235 5000
; TELEFAX: (201) 235 3500
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Actinomyces viscosus/SOD gene
; US-08-133-711-44

Query Match      8.4%; Score 48; DB 1; Length 491;
Best Local Similarity 50.9%; Pred. No. 0.13;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 147 GCCCGCGCGCTTGGCGCGCCAGCGCGCGCTCCAGCAGCAGCTGCTCAAGAGAAG 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 GCGCTGGAGGCTTGGCGCGCGCGCGCGAGGAGCGGACCTGGGTGCATCAACCTGTGG 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 TTCGAGGAGTGCACAGCTTCAACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 GAGAGAACCTTCAACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 CAGCGCGCGCGCGCTTGGTGGCGGAGACCGAGCGTTCCTCGCCAGGCG 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 CTCTCCCGAACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 ACCGCGGAGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 TTCGGCTCTTGGAGAAGTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 274

RESULT 13
US-08-440-856A-2
; Sequence 2, Application US/08440856A
; Patent No. 5750873
; GENERAL INFORMATION:
; APPLICANT: DELLAPORTA, STEPHEN L.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
; TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE. N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,856A
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A.
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 05463-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1517
; TELEFAX: (202) 887-0763
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1187 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-440-856A-2

Query Match      8.4%; Score 48; DB 1; Length 1187;
Best Local Similarity 47.6%; Pred. No. 0.14;
Matches 141; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
```

```
QY 79 ACCAAGCGTCGGCACCAATGCTTACTACCAGAGGTGGACTACTGCTCGGAGAGGTGA 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 AGCAGCGCGCTCGCCATGACCCAGCGCCCGCGGAGCATCATCTCCGTGCGCCAGCG 594
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 139 GGTGGTGGCCCCCGCGGCTTCGGCGCGCCACGCGCGGCGTCCAGCAGCAGCAGCTGCTCA 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 TCGCGCGCGTCTCGCGCGGCTTCGGCGCGCCACGCGCTACACGCGCTCCCAAGCAGCGCATCG 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 AGGAGAAGTTCGAGGAGTTCACACGCTCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCG 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 TGGGGCTACCAAGAACGCGCGCTTCGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 ATGTCACACGCGCGCGCGCTTCGTGGTGGCGGAGACGAGGTGGAAGAGGACATCA 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 TCTCCCTTCGGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 319 ACACCTGACCGCGCGGTCACGAGCGGAGGAGAGCTTCTCGCGCGCGCGCGCGCGCGCG 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 CCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 830

RESULT 14
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Query Match      8.3%; Score 47.8; DB 4; Length 4403765;
Best Local Similarity 49.8%; Pred. No. 0.34;
Matches 121; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 151 CGCGCGGCTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3921743 CGCGCGCGCGCGCAACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 AGGAGTTCGACACGCTCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3921803 GCGGGATCGCGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921862
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 GCGGCGACGCTTCGTGGTGGCGGAGACGAGGTGGAAGAGGACATCAACACCTGCGCGCG 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3921863 CCGGCAACACAGGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921922
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 GCGAGTTCACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3921923 GAGGCGTGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921982
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 391 CGG 393
   |||
Db 3921983 CGG 3921985
   |||
```

Mon Jun 16 14:55:57 2003

```
RESULT 15
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
; FILE REFERENCE: 0867/OD905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4
```

```
Query Match      8.2%; Score 47.2; DB 4; Length 1926;
Best Local Similarity 49.6%; Pred. No. 0.21;
Matches 121; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY      109 AGGAGGTGACTACTGCTCGAGGAGGTGAGTGGTGGCCCGCGGCTTCGGCCGCC 168
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      924 ACGGGAGGACGAGGACGGGGAGGACGGGGAGGACGAGGACGGGGAGGACG 865

QY      169 ACGGGCGCGCGTCCAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGTTCGACACGGTCT 228
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      864 AGGACGGGGAGGACGGGGAGGAGGACGAGGACGGGGAGGACGAGGAGGACG 805

QY      229 CACGCGCGCGCGCCAAACCCACCACCATGGTCCACCGCGCGGCCACGGCTTCGTGG 288
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      804 GGGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGCGGG 745

QY      289 TCGCGGAGACCGGGTCCGAGAGGACATCAACACCTGCACCGCGCGAGGTCCACGAGCGCA 348
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      744 AGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGCGGG 585

QY      349 GGGA 352
      |||
Db      584 AGGA 581
```

Search completed: June 16, 2003, 03:58:13
Job time : 81.4096 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 23:41:35 ; Search time 188.272 Seconds
(without alignments)
5865.849 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

ID	Seq	Score	Match	Length	ID	Description
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*	62.6	10.9	390	13	AAQ21833
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*	62.6	10.9	390	14	AAQ36859
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*	62.6	10.9	390	22	AAQ76910
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*	62.6	10.9	390	24	AAI72775
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*	53.2	9.3	4403765	22	AAI99683
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*	53.2	9.3	4411529	22	AAI99682
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*	52.8	9.2	1128	20	AAI10560
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*	52.8	9.2	1128	21	AAI10560
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*	52.8	9.2	1128	21	AAA46026

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.6	10.9	390	13	AAQ21833
2	62.6	10.9	390	14	AAQ36859
3	62.6	10.9	390	22	AAQ76910
4	62.6	10.9	390	24	AAI72775
5	53.2	9.3	4403765	22	AAI99683
6	53.2	9.3	4411529	22	AAI99682
7	52.8	9.2	1128	20	AAI10560
8	52.8	9.2	1128	21	AAI10560
9	52.8	9.2	1128	21	AAA46026

C 10	52.8	9.2	1128	24	AAQ27892	Human G-protein co
C 11	52.8	9.2	1128	24	AAQ27892	Human DNA for pote
C 12	52.8	9.2	109519	22	AAQ80649	Micromonospora DNA
C 13	52.8	9.2	114955	20	AAQ80693	Human adenosine A1
C 14	51.8	9.0	2244	17	AAQ53491	Mature isoenzyme A1
C 15	51.8	9.0	2244	19	AAQ10429	Flavobacterium iso
C 16	51.8	9.0	2334	17	AAQ10428	isoamylase gene.
C 17	51.8	9.0	2334	19	AAQ23639	Full length Flavob
C 18	51.4	9.0	3849	22	AAQ25795	S. chrysomallus ac
C 19	51	8.9	4411529	22	AAI99682	Mycobacterium tube
C 20	50.8	8.9	3059	8	AAQ70566	Sequence encoding
C 21	49.8	8.7	522	21	AAQ248813	Wheat inositol 1.3
C 22	49.6	8.6	685	24	ABQ78045	Wheat SCIP-1 ortho
C 23	49.4	8.6	65140	22	AAQ17184	Streptomyces nous
C 24	49.4	8.6	125401	22	AAQ17186	Streptomyces nous
C 25	49.2	8.6	1171	22	AAQ44082	Oryza sativa perox
C 26	49.2	8.6	6781	19	AAQ45824	Maize phosphoenolp
C 27	49.2	8.6	7559	20	AAQ82011	Pseudomonas aerugi
C 28	49	8.5	1470	23	AAQ54218	Maize T82 CDNA nuc
C 29	48.6	8.5	1236	16	AAQ54060	Streptomyces nous
C 30	48.2	8.4	27541	22	AAQ17185	Superoxide dismuta
C 31	48.2	8.4	125401	22	AAQ17186	Trehalose-releasin
C 32	48	8.4	490	15	AAQ62127	Streptomyces nous
C 33	48	8.4	1725	21	AAQ10503	Streptomyces nous
C 34	48	8.4	2218	21	AAQ10516	Trehalose-releasin
C 35	47.8	8.3	4403765	22	AAI99683	Mycobacterium tube
C 36	47.4	8.3	14806	22	ABQ3809	Streptomyces galli
C 37	47.2	8.2	1464	24	ABQ90041	M. capsulatus gene
C 38	47	8.2	1260	11	AAQ05666	TGF-57-Pseudomonas
C 39	47	8.2	1260	11	AAQ06127	Sequence encoding
C 40	47	8.2	1260	18	AAQ72116	TGF-57-PE40 gene d
C 41	47	8.2	1260	20	AAQ76879	Wild type TGF- α
C 42	47	8.2	2291	17	AAQ16859	Nitric oxide synth
C 43	47	8.2	2668	21	AAQ00337	Wheat raffinose sy
C 44	47	8.2	3616	17	AAQ16858	Nitric oxide synth
C 45	47	8.2	4089	17	AAQ16857	Bovine endothelial

ALIGNMENTS

RESULT 1
AAQ21833
ID AAQ21833 standard; DNA; 390 BP.
AC AAQ21833;
XX
XX
DT 08-JUN-1992 (first entry)
XX
DE Randomising oligonucleotide used in SPERT mRNA prepn.
DE Systematic polypeptide evolution by reverse translation; SPERT;
KW ligand binding; ss.
XX
OS Synthetic.
XX
PN WO9202536-A.
XX
PD 20-FEB-1992.
XX
PF 01-AUG-1991; 91WO-US05463.
XX
PR 02-AUG-1990; 90US-0561968.
XX
XX (COLS) UNIV OF COLORADO.
XX Gold L, Tuerk C;
XX WPI; 1992-080018/10.
XX New method of systematic polypeptide evolution by reverse
PT translation - by linking each polypeptide in sample mixt. to
PT individualised mRNA allowing further synthesis of selected

PT	polypeptide(s)	
XX	Example; Page 55; 102pp; English.	
PS		
CC	The sequence is that of an example randomising oligonucleotide which	
CC	is used in the prepn. of mRNA encoding candidate polypeptides for the	
CC	method of systematic polypeptide evolution by reverse translation	
CC	(SPERT). The method provides a rapid way of isolating and identifying	
CC	polypeptide ligands which bind to target mols. The polypeptide ligands	
CC	can be used in e.g. assay methods, diagnostic procedures, cell sorting,	
CC	as activators or inhibitors of target mol. function, as probes, as	
CC	sequestering agents, drug delivery vehicles, modifiers of hormone	
CC	action and as catalysts. See also AAQ21830-Q21832.	
XX		
SQ	Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;	
	Query Match 10.9%; Score 62.6; DB 13; Length 390;	
	Best Local Similarity 52.0%; Pred. No. 0.0016;	
	Matches 140; Conservative 0; Mismatches 129; Indels 0; Gaps 0;	
QY	103 ACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGCGGCTTCG 162	
DB	109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168	
QY	163 GCCGCCACGCGCGGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAGTTCGACA 222	
DB	169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228	
QY	223 CGGTCTCACGCGCGCGGCCAACCCACGACCATGTCACACCGCGGCGGCT 282	
DB	229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288	
QY	283 TCGTGGTGGCGAGACGAGGTGCGAGAGGACATCAACACCTGCACCGCGGAGTCCACG 342	
DB	289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348	
QY	343 ACGCAGGGAGAGCTTCTCGCCAGGGCT 371	
DB	349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377	
	RESULT 2	
AAQ36859		
ID	AAQ36859 standard; DNA; 390 BP.	
XX		
AC	AAQ36859;	
XX		
DT	22-JUN-1993 (first entry)	
XX		
DE	PCR primer for 5' fixed sequence contg. T7 promoter and RBS.	
XX		
KW	Systematic peptide evolution by reverse translation; SPERT; ligand;	
KW	specific; inhibitors; probes; assay; cell sorting; ss.	
OS	Synthetic.	
XX		
PN	WO9303172-A.	
XX		
PD	18-FEB-1993.	
XX		
PF	31-JAN-1992; 92WO-00801.	
XX		
PR	01-AUG-1991; 91US-0739055.	
XX		
PA	(UYRE-) UNIV RES CORP.	
XX		
PI	Gold L, Pribnow D, Smith JD, Tuerk C;	
XX		
DR	WPI; 1993-076529/09.	
XX		
PT	Systematic polypeptide evolution by reverse translation - used	
PT	for prodn. of polypeptide ligand specific for desired target	
PT	molecule	

XX	Example 1; Page 84; 98pp; English.	
PS		
CC	SPERT is used to select novel polypeptides that bind the antibody	
CC	of the epitope commonly recognised by the antiserum from autoimmune	
CC	mice which are the fl progeny of a cross of NZB and NZW parents	
CC	(Portanova et al., J. Immunol. 144, 4633, 1990). The known epitope	
CC	consists of ca. 10 amino acids at the N-terminus of the histone H2B	
CC	protein. To make mRNA encoding candidate polypeptides a 5' fixed	
CC	sequence composed of a T7 promoter sequence and a ribosome binding	
CC	site which is recognised by both prokaryotic and eukaryotic ribosomes,	
CC	terminating in a restriction endonuclease site is synthesised and cloned	
CC	using a number of oligonucleotides (example shown). A 3' fixed sequence	
CC	is placed into a restriction site to provide an mRNA encoding the C-	
CC	terminal trailer sequence of ca. 100 nucleotides lacking stop codons.	
CC	In addition, a 3' primer annealing site is provided so that cDNA	
CC	synthesis can be accomplished on the mRNA recovered from partitioned	
CC	ribosome complexes. See also AAQ36845-63.	
XX		
SQ	Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;	
	Query Match 10.9%; Score 62.6; DB 14; Length 390;	
	Best Local Similarity 52.0%; Pred. No. 0.0016;	
	Matches 140; Conservative 0; Mismatches 129; Indels 0; Gaps 0;	
QY	103 ACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGCGGCTTCG 162	
DB	109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168	
QY	163 GCCGCCACGCGCGGCTCCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAGTTCGACA 222	
DB	169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228	
QY	223 CGGTCTCACGCGCGCGGCCAACCCACGACCATGTCACACCGCGGCGGCT 282	
DB	229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288	
QY	283 TCGTGGTGGCGAGACGAGGTGCGAGAGGACATCAACACCTGCACCGCGGAGTCCACG 342	
DB	289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348	
QY	343 ACGCAGGGAGAGCTTCTCGCCAGGGCT 371	
DB	349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377	
	RESULT 3.	
AAAF76910		
ID	AAAF76910 standard; DNA; 390 BP.	
XX		
AC	AAAF76910;	
XX		
DT	29-MAY-2001 (first entry)	
XX		
DE	Sequence containing a 120 repeat of ACG flanked by fixed fragments.	
XX		
KW	Ligand isolation; systemic polypeptide evolution by reverse translation;	
KW	SPERT; ss.	
OS	Synthetic.	
XX		
PN	US6194550-B1.	
XX		
PD	27-FEB-2001.	
XX		
PF	23-NOV-1998; 98US-0197649.	
XX		
PR	31-JAN-1992; 92US-0829461.	
XX		
PR	02-AUG-1990; 90US-0561968.	
XX		
PR	01-AUG-1991; 91US-0739055.	
XX		
PA	(GOLD/) GOLD L.	
PA	(TUERK/) TUERK C.	


```

103 ACTACGAGGAGTGGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGCGGCTTCG 162
109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168
163 GCGCGCAGCGCGCGCGCTTCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAGTTCG 222
169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228
223 CGTCTCTACGCGCGCGCGCAACACCAACCAACCAACCAACCAACCAACCAACCA 282
229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288
283 TCGTGTGCGCGAGACGAGGTCGAAGAGGACATCAACCTGCGACCGCGGAGTTCG 342
289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348
343 ACGCGAGGAGGAGTTCCTCGCCAGGGCT 371
349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377

RESULT 5
AAI99683/C
ID AAI99683 standard; DNA; 4403765 BP.
XX AC
XX AA199683;
XX
DT 15-JAN-2002 (first entry)
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ
XX
PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
```

```

Query Match 9.3%; Score 53.2; DB 22; Length 4403765;
Best Local Similarity 53.3%; Pred. No. 0.28;
Matches 112; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 147 GCGCGCGCGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 206
DB 1191569 GCGCGCGCGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191510
QY 207 TTCAGGAGGTTCGACGCGTCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 266
DB 1191509 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191450
QY 267 CACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 326
DB 1191449 GACGCGCGGTGCGCGCGGTGCTGTCGCGCAACGCGCGCGCGCGCGCGCGCG 1191390
QY 327 ACCGCGCGAGGTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 356
DB 1191389 GCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191360

RESULT 6
AAI99682
ID AAI99682 standard; DNA; 4411529 BP.
XX AC
XX AA199682;
XX
DT 15-JAN-2002 (first entry)
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
DR WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ
XX
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
```


PA (AREN-) ARENA PHARM INC.
 XX Chen R, Dang HT, Liaw CW, Lin I;
 PI WPI: 2000-400068/34.
 XX P-PSDB; AAV71298.
 DR Novel human orphan G protein-coupled receptors and the encoding cDNAs
 XX for use in the identification of G protein-coupled receptor agonists -
 PT Claim 29; Page 61-62; 102pp; English.
 XX
 XX The present sequence is a cDNA encoding hGPCR27, an endogenous human
 CC orphan G protein-coupled receptor (GPCR), expressed in foetal brain,
 CC putamen, pituitary gland and caudate nucleus. The hGPCR27 cDNA was
 CC identified using EST (expressed sequence tag) AA75870 as a probe.
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
 CC alpha helices with an extracellular N-terminus and an intracellular
 CC C-terminus. However, no endogenous ligands has yet been identified for
 CC the proteins of the invention. The orphan GPCRs may be used in the
 CC identification of their endogenous ligands, and to screen potential GPCR
 CC agonists and antagonists for use as pharmaceutical agents. The proteins
 CC may also be used in the study of GPCR-mediated signalling cascades, and
 CC to elucidate their precise role in normal and diseased human conditions.
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue
 CC localisation expression analysis to provide information about their
 CC function in healthy and pathological states.
 XX
 XX Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;
 SQ
 Query Match 9.2%; Score 52.8; DB 21; Length 1128;
 Best Local Similarity 52.7%; Pred. No. 0.15;
 Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;
 QY 97 TGGCTTACTACAGGAGGTGGACTACTGCTCGAGAGGTGAGTGGTGGCCCGCGCG 156
 DB 721 TGGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 662
 QY 157 GCTTTCG 216
 DB 661 GCACAGGCG 603
 QY 217 TCGACAGGTCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
 DB 602 TAGACAGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 543
 QY 277 ACAGGCTTCGTGGTGGCGAGACCGAGGTCGGAAGAGGACATCAACACTGACCGCGGAGG 336
 DB 542 GGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 QY 337 TCCACGAGCGCGAGGAGCG 356
 DB 482 TCCAGCAGTGGCGGGAAGGC 463
 RESULT 9
 AAA46026/c
 ID AAA46026 standard; cDNA; 1128 BP.
 XX
 XX AAA46026;
 AC
 XX
 XX 22-AUG-2000 (first entry)
 DE Human G protein coupled receptor hGPCR27 encoding cDNA SEQ ID NO:17.
 XX
 XX Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200022131-A2.
 PN
 XX

PD 20-APR-2000.
 XX 13-OCT-1999; 99WO-US24065.
 PF 13-OCT-1998; 98US-0170496.
 XX 12-NOV-1998; 98US-0108029.
 PR 27-NOV-1998; 98US-0109213.
 PR 16-FEB-1999; 98US-0110060.
 PR 26-FEB-1999; 98US-0120416.
 PR 12-MAR-1999; 98US-0121852.
 PR 12-MAR-1999; 98US-0123944.
 PR 12-MAR-1999; 98US-0123945.
 PR 12-MAR-1999; 98US-0123946.
 PR 12-MAR-1999; 98US-0123948.
 PR 12-MAR-1999; 98US-0123951.
 PR 28-MAY-1999; 98US-0136436.
 PR 28-MAY-1999; 98US-0136437.
 PR 28-MAY-1999; 98US-0136439.
 PR 28-MAY-1999; 98US-0137127.
 PR 28-MAY-1999; 98US-0137131.
 PR 28-MAY-1999; 98US-0137567.
 PR 27-AUG-1999; 98US-0141448.
 PR 03-SEP-1999; 98US-0151114.
 PR 29-SEP-1999; 98US-0152524.
 PR 29-SEP-1999; 98US-0156633.
 PR 29-SEP-1999; 98US-0156555.
 PR 29-SEP-1999; 98US-0156634.
 XX
 XX (AREN-) ARENA PHARM INC.
 XX Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX WPI: 2000-317986/27.
 DR P-PSDB; AAB02832.
 DR
 XX Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 1; Page 91-92; 187pp; English.
 XX The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 XX Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;
 SQ
 Query Match 9.2%; Score 52.8; DB 21; Length 1128;
 Best Local Similarity 52.7%; Pred. No. 0.15;
 Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;
 QY 97 TGGCTTACTACAGGAGGTGGACTACTGCTCGAGAGGTGAGTGGTGGCCCGCGCG 156
 DB 721 TGGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 662
 QY 157 GCTTTCG 216
 DB 661 GCACAGGCG 603
 QY 217 TCGACAGGTCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
 DB 602 TAGACAGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 543
 QY 277 ACAGGCTTCGTGGTGGCGAGACCGAGGTCGGAAGAGGACATCAACACTGACCGCGGAGG 336
 DB 542 GGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483

QY 337 TCACGAGCGCAGGAGACG 356
 DB 482 TCACGACTGGCGGGAAGC 463

RESULT 10

AAD27892/C
 ID AAD27892 standard; DNA; 1128 BP.

XX AAD27892;

DT 31-MAY-2002 (first entry)

DE Human G-protein coupled receptor 14266 DNA.

KW Human: G-protein coupled receptor; GPCR 14266; haematopoietic disorder;
 KW neutrophil deficiency disorder; splenomegaly; pulmonary hypertension;
 KW colon; diarrhoea; hepatic injury; idiopathic inflammatory bowel disease;
 KW uterus; endometriosis; brain; acute meningitis; multiple sclerosis;
 KW T-cell; systemic lupus erythematosus; skin; vitiligo; heart failure;
 KW angina pectoris; atherosclerosis; haemolytic anaemia; thymic cyst; ds.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..1128

FT /*tag= a

FT /product= "GPCR 14266"

FT sig_peptide 1..102

FT /*tag= b

FT mat_peptide 103..1125

FT /*tag= c

FT /product= "Mature GPCR 14266"

XX WO200212344-A2.

XX 14-FEB-2002.

XX 08-AUG-2001; 2001WO-US24835.

XX 09-AUG-2000; 2000US-0634392.

XX (MILL-) MILLENNIUM PHARM INC.

XX Welch NS;

XX WPI; 2002-217181/27.

XX P-PSDB; AAE14597.

XX Identifying compound that binds to and/or modulate activity of the
 PT polypeptide for treating splenomegaly, emphysema and multiple
 PT sclerosis, comprises using human G-protein coupled receptor 14266
 PT polypeptide -

XX Claim 1; Fig 2; 115pp; English.

XX The invention relates to a method of identifying an agent that binds to
 CC and/or modulates activity of human G-protein coupled receptor (GPCR)
 CC 14266. The invention also provides a method for modulating the level
 CC of activity of the GPCR in cells in vivo or in vitro, or in a subject
 CC predisposed to having a haematopoietic or neutrophil deficiency
 CC disorders. Detecting 14266 polypeptides or nucleic acids in specific
 CC biological samples is useful in disease diagnosis. Modulating level or
 CC activity of the GPCR polypeptide or polynucleotide is useful for treating
 CC diseases related to 14266 receptor malfunction e.g., disorders involving
 CC spleen (e.g. splenomegaly, neoplasms); lung (e.g. pulmonary hypertension,
 CC emphysema); colon (e.g. diarrhoea and dysentery, idiopathic inflammatory
 CC bowel disease); liver (e.g. hepatic injury, viral hepatitis); uterus and
 CC endometrium (e.g. endometriosis, endometrial polyps); brain (e.g.
 CC acute meningitis, multiple sclerosis, spinocerebellar degeneration);
 CC T-cells (e.g. systemic lupus erythematosus, polyarteritis nodosa); skin
 CC (e.g. vitiligo, acanthosis nigricans); heart (e.g. heart failure,

CC angina pectoris); blood vessels (e.g. atherosclerosis, Raynaud disease,
 CC thalassaemia syndromes, haemolytic anaemia) and disorders of the thymus
 CC (e.g. thymic cysts, Hodgkin's disease). The present sequence is human
 CC GPCR 14266 DNA.

SQ Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;

Query Match 9.2%; Score 52.8; DB 24; Length 1128;
 Best Local Similarity 52.7%; Pred. No. 0.15;
 Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

QY 97 TGGCTTACTACGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGCGG 156
 DB 721 TGGCGCGCGCTGGCGCGTGGCGCGCGCGCGCGCGTGGAGGTCCAGTCTGTGCTGACGCGG 662
 QY 157 GCTTTCGCGCGCACGCGCGCGGTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAGG 216
 DB 661 GCACGAGCGCGCGCGCGCGCATCTTGGCGGCGGTCGTGGATGAAGAGCAGCGCGGAGG 603
 QY 217 TCACACGGTCTCACGCGCGCGCGCGCCCAACACACACACCATGTTGTCACCGCGCGG 276
 DB 602 TAGACGAGGTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 543
 QY 277 ACGGCTTCGTGGTGGCGGAGCAGCGGTGCAAGAGGACATCAACACCTGCACCGCGGAGG 336
 DB 542 GGGCGCGCGTGGCGCGCGCTCTCCAGGCGCGCGCGCGTCTGCTGCTGCGCACGCGCGG 483
 QY 337 TCACGAGCGCAGGAGAGC 356
 DB 482 TCACGACTGGCGGGAAGGC 463

RESULT 11

AAS98049/c

ID AAS98049 standard; DNA; 1128 BP.

XX AC AAS98049;

XX 12-MAR-2002 (first entry)

XX Human DNA for potential G protein-coupled receptor #7.

XX Human: G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
 KW Alzheimer's disease; amyotrophic lateral sclerosis; aschma;
 KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
 KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
 KW depression; epilepsy; macular degeneration; lymphoma; melanoma;
 KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
 KW psoriasis; rheumatoid arthritis; schizoprenia; ulcerative colitis;
 KW tuberculosis; cognition disorder; memory disorder; anorexia;
 KW hormonal release disorder; cardiovascular activity disorder;
 KW pain perception disorder; obesity; diabetes; obesity;
 KW diabetes; hyperlipidaemia; stroke; gene therapy.

XX Homo sapiens.

XX WO200185791-A1.

XX 15-NOV-2001.

XX 11-MAY-2001; 2001WO-US15332.

XX 11-MAY-2000; 2000US-203217P.

XX 18-MAY-2000; 2000US-205945P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;

XX WPI; 2002-066595/09.

XX Novel G protein-coupled receptor polypeptides including galanin
 PT receptor polypeptides useful for identifying modulators that are useful

PT for treating Alzheimer's disease, psoriasis, melanoma, multiple
PT sclerosis, stroke
XX
PS
XX
XX
CC The invention relates to an isolated polypeptide encoded by a
CC nucleic acid molecule that is at least 80% identical to the G
CC protein-coupled (GPCR) polynucleotides included in the specification.
CC Also included are probes based on the GPCR sequences (including
CC antisense probes), a host cell comprising an expression vector comprising
CC the GPCR sequence, antibodies raised against the polypeptides,
CC and methods of identifying modulators of the polypeptides. The
CC polypeptides are useful for identifying modulator compounds which
CC function as modulators, activators, repressors, agonists or antagonists
CC of the novel GPCR polypeptides including the G4U polypeptide. The
CC antibodies and nucleic acid probes as described above can be used to
CC detect the presence of the polypeptides and nucleic acids and are used to
CC diagnose a variety of diseases or disorders in which GPCRs are involved
CC e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,
CC atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,
CC chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,
CC depression, epilepsy, macular degeneration, lymphoma, melanoma,
CC multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,
CC psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,
CC tuberculosis and many other diseases listed in the specification. The
CC probes and antibodies are also useful for diagnosing cognition and memory
CC disorders, anorexia, hormonal release disorders, cardiovascular activity
CC disorders, pain perception disorders, obesity, diabetes, Alzheimer's
CC disease. Preferably, compounds that decrease or increase
CC the expression of galanin receptor (GAL4) can be used to treat obesity,
CC diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is
CC useful for treating the above mentioned disorders by gene therapy
CC techniques. The present sequence is a novel GPCR polynucleotide of the
CC invention.
XX
SQ Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;
Query Match 9.2%; Score 52.8; DB 24; Length 1128;
Best Local Similarity 52.7%; Pred. No. 0.15;
Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;
QY 97 TGCTTACTACAGGAGGTGGACTACTGTCGAGGAGGTGAGTGGTGGCCCGCGG 156
Db 721 TGGCGCGCGCTGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 562
QY 157 GCTTCGG 216
Db 661 GCACGAGGCGG 603
QY 217 TCGACACGCGTCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 276
Db 602 TAGACGAGGTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 543
QY 277 ACGGCTTCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 336
Db 542 GGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 483
QY 337 TCCACGAGCGG 356
Db 482 TCCAGCAGTGGCGGGAAGGC 463
RESULT 12
AAS08693
ID AAS08693 standard; DNA; 109519 BP.
XX AAS08693;
AC AAS08693;
XX
XX 26-SEP-2001 (first entry)
XX Micromonospora DNA encoding biosynthetic enzymes for Everninomycin.
DE
DE
DE
KW Everninomycin; antibiotic; bottle-neck gene; orthomycin;

fermentation; ds.
KW XX OS Micromonospora carbonacea var. africana.
XX
XX
FH Key Location/Qualifiers
FT CDS complement (132..1382)
FT FT
FT *tag= a
FT /product= "EvdA"
FT complement (1389..1394)
FT *tag= b
FT complement (1490..2611)
FT *tag= c
FT /product= "EvdB"
FT complement (2618..2622)
FT *tag= d
FT complement (2622..3860)
FT *tag= e
FT /product= "EvdC"
FT complement (3867..3870)
FT *tag= f
FT 4143..5312
FT *tag= g
FT /product= "EvdD"
FT 4134..4138
FT *tag= h
FT 5309..6235
FT *tag= i
FT /product= "EvdE"
FT 6232..7275
FT *tag= j
FT /product= "EvdF"
FT 6226..6229
FT *tag= k
FT 7272..8327
FT *tag= l
FT /product= "EvdG"
FT 8342..9364
FT *tag= m
FT /product= "EvdH"
FT 8333..8336
FT *tag= n
FT complement (9463..10224)
FT /product= "EvdI"
FT complement (10232..10235)
FT *tag= o
FT 10424..11176
FT *tag= q
FT /product= "EvdJ"
FT 12027..12455
FT *tag= r
FT /product= "EvdK"
FT partial
FT /note= "No start codon"
FT complement (12108..13022)
FT *tag= s
FT /product= "EvdL"
FT complement (13027..13030)
FT *tag= t
FT complement (14410..15363)
FT *tag= u
FT /product= "EvdR"
FT complement (15369..15373)
FT *tag= v
FT complement (15380..16414)
FT *tag= w
FT /product= "EvdS"
FT complement 16419..17873
FT *tag= x
FT /product= "EvdT"
FT complement (17870..18934)
FT *tag= y
FT /product= "EvdU"

RESULT 13

AA53491/c
ID AAX53491 standard; DNA; 114955 BP.XX
XX
AC AAX53491;XX
XX
DT 05-JUL-1999 (first entry)XX
XX
DE Human adenosine A1 receptor antisense oligonucleotide fragment.XX Antisense oligonucleotide; multiple target; antisense treatment;
XX impaired respiration; inflammation; lung disease;
XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
XX acute asthma; allergy; asthma; impeded respiration;
XX respiratory distress syndrome; pain; cystic fibrosis;
XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
XX prostate cancer; ss.XX
OS Synthetic.XX
XX WO9913886-A1.XX
XX PD 25-MAR-1999.XX
XX PF 17-SEP-1998; 98WO-US19419.XX
XX PR 09-JUN-1998; 98US-0093972.XX
XX PR 17-SEP-1997; 97US-0059160.XX
XX PA (UYEC-) UNIV EAST CAROLINA.XX
XX PI Nyce JW;XX
XX DR WPI; 1999-229400/19.XX
XX PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
PT vasoconstrictionXX
XX PS Disclosure; Page 37; 120pp; English.XX
XX CC The specification describes antisense oligonucleotides (AAX52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, gene
CC initiation codons, genomic flanking regions, intron-exon borders, the
CC 5'-end, the 3'-end and the juxta-section between coding and non-coding
CC regions and all segments of RNAs encoding proteins associated with one
CC or more diseases, conditions or mixtures. The antisense oligonucleotides
CC may be derived from sequences AAX5272-74. These multiple target
CC oligonucleotides (specifically AAX55180-271) can be used for the
CC antisense treatment of diseases and conditions. Typical diseases and
CC conditions are those associated with impaired respiration and
CC inflammation, including lung diseases, pulmonary vasoconstriction,
CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
CC hepatic metastases, as well as all types of cancers which may metastasize
CC or have metastasized to the lungs, including breast and prostate cancer.XX
XX SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;XX
XX Query Match 9.2%; Score 52.8; DB 20; Length 114955;XX
XX Best Local Similarity 35.6%; Pred. No. 0.24;XX
XX Matches 144; Conservative 29; Mismatches 232; Indels 0; Gaps 0;XX
XX 1 ACCACGGCTCCCGCCACGGCTCCGACAGCAATCCACACAGCATTCGACGTCACAG 60

Db 103487 ACGGCGCGCNCNNNNNSGCCGCAVAGACGGCGCGCNCNNNNNSGCCGCAVAGACGGCGG 103428
Qy 61 GGCGCTGGCAGACACACCAAGCGTGGGACCAATGCTTACTACCAAGAGGTGGACT 120
Db 103427 CGCNCNNNNNSACGGCGCCCAVAGACGGCGCGCNCNNNNNSGACGCCCAVAGACGGCGCGC 103368
Qy 121 ACTGCTCGGAGGAGGTGAGGTGCGTGGCGCGCGCGCTTCGGCGCGCCACGCGCGCGCGG 180
Db 103367 NNNNNNSGACGCCCGCAVAGACGGCGCGCNCNNNNNSCCGAGCGCGCCCAVAGACGGCGCGG 103308
Qy 181 TCCACGACAGCTGCTCAAGGAGAGATTCTGAGGAGGTGACACAGGTCTCACGCGCGCGCGG 240
Db 103307 CNCNNNNNSCCGACGCCCAVAGACGGCGCGCNCNNNNNSGCCGCGCCCAVAGACGG 103248
Qy 241 CCAACACACACACACCATGTGTACACGCGCGCGCGCTTCGTGTGCGCGGAGACCA 300
Db 103247 CGCGCNCNNNNNSGCCGCGCCCAVAGACGGCGCGCNCNNNNNSGCCGCGCGCGG 103188
Qy 301 GGGTCGAGAGGACATCAACACCTGCACGCGCGAGGTCCACGAGCGCGAGGAGAGCTTCC 360
Db 103187 CCAVAGACGGCGCGCNCNNNNNSCCGCGCGCGCAVAGACGGCGCGCNCNNNNNSCC 103128
Qy 361 TCGCCAGGCTAACTGAGCGCGCGCGCGCGCATCCACGCGCG 405
Db 103127 CGGCGCGAGCGCGCCCAVAGACGGCGCGCNCNNNNNSGCCGCGCGCGG 103083

RESULT 14

AAT10429

ID AAT10429 standard; DNA; 2244 BP.

XX
XX AC AAT10429;XX
XX DT 03-JUN-1996 (first entry)XX
XX DE Mature isoamylase DNA.XX
XX KW Isoamylase; starch; amylose; amylopectin; transgenic plant;
XX fructose; ds.XX
XX OS Flavobacterium sp.XX
XX PN WO9603513-A2.XX
XX PD 08-FEB-1996.XX
XX PF 24-JUL-1995; 95WO-US09323.XX
XX PR 28-JUL-1994; 94US-0281902.XX
XX PA (MONS) MONSANTO CO.XX
XX PI Barry GF, Kishore GM, Krohn BM;XX
XX DR WPI; 1996-117056/12.XX
XX DR P-PSDB; AAR92693.XX
XX PT New isolated DNA encoding isoamylase - used to produce plants having
PT a modified starch content, e.g. a higher ratio of amylose to
PT amylopectinXX
XX PS Claim 2; Page 35-39; 46pp; English.XX
XX CC A DNA sequence (AAT10429) codes for the mature isoamylase (plus an
XX N-terminal methionine) of Flavobacterium sp. (see also AAR92692).
XX It was obtd. by amplification of the cloned isoamylase gene
XX (AAT10428) using primers (AAT10432-33) designed to add a start codon
XX to the beginning of the processed mature enzyme (Ala-33) and a
XX stop codon adjacent to the endogenous TGA codon. The DNA sequence
XX is used for the recombinant prodn. of isoamylase or can be
XX expressed in transgenic plants, e.g. potato, cassava, sweet potato,
XX corn, wheat, barley or rice, to alter the amylose:amylopectin

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run On:      June 15, 2003, 23:43:51 ; Search time 1684.35 seconds
              (without alignments)
              9917.805 Million cell updates/sec

Title:       US-10-090-035-3
Perfect score: 574
Sequence:    1 accacagatgccaccacg
```

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_lm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hgto_hum.*
40: em_hgto_mus.*
41: em_hgto_other.*

Pred. No. Is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
C 1	174.6	30.4	173699	8	AP003416	Oryza sativa
C 2	174.6	30.4	174218	2	AC091071	Oryza sativa
C 3	174.6	30.4	185095	8	AC003436	Oryza sativa
C 4	174.6	30.4	185481	2	AC091088	Oryza sativa
C 5	143.8	25.1	562	8	SST242802	Sporobolus
C 6	62.6	10.9	390	6	AR135147	Sequence
C 7	61.4	10.7	35100	1	SC3A7	AL031155 Streptomyces
C 8	61.4	10.7	150676	2	AC107225	Oryza sativa
C 9	58.2	10.1	29080	1	SCBNC2861	AL593842 Streptomyces
C 10	57.8	10.1	18420	1	SCG30	AL352972 Streptomyces
C 11	57	9.9	4851	1	SCOI18817	Y18817 Streptomyces
C 12	56.8	9.9	150053	8	AP003735	Oryza sativa
C 13	56.2	9.8	102689	2	AP004045	Oryza sativa
C 14	56.2	9.8	138289	2	AP004873	Oryza sativa
C 15	54.6	9.5	3164	1	AF263463	Oryza sativa
C 16	54.4	9.5	13368	1	AE263463	Streptomyces
C 17	54.2	9.4	106050	2	AE005107	Halobacterium
C 18	54.2	9.4	146436	2	AP004088	Oryza sativa
C 19	54.2	9.4	152263	2	AC108759	Oryza sativa
C 20	54.2	9.4	187705	2	AP005412	Oryza sativa
C 21	54	9.4	36368	1	SC9B5	Oryza sativa
C 22	53.8	9.4	174410	2	AC128158	Streptomyces
C 23	53.6	9.3	34593	1	SC2A8	AL028158 Rattus norvegicus
C 24	53.4	9.3	10612	1	AF086832	Streptomyces
C 25	53.4	9.3	41173	1	SC141	AL589164 Streptomyces
C 26	53.4	9.3	65547	8	AY013245	Streptomyces
C 27	53.4	9.3	120159	2	AC120509	Streptomyces
C 28	53.2	9.3	15041	1	AE006391	Oryza sativa
C 29	53.2	9.3	47852	1	MTV023	Oryza sativa
C 30	53	9.2	136150	8	AP002485	Oryza sativa
C 31	52.8	9.2	1128	6	AX393069	Mycobacterium
C 32	52.8	9.2	1128	9	AB040799	Oryza sativa
C 33	52.8	9.2	1245	6	AC1195950	Sequence
C 34	52.8	9.2	109519	6	AC1195929	Homo sapiens
C 35	52.8	9.2	163494	2	AC016964	Sequence
C 36	52.8	9.2	171397	9	AC096970	Oryza sativa
C 37	52.8	9.2	190506	2	AC068990	Homo sapiens
C 38	52.6	9.2	179198	2	AC130729	Oryza sativa
C 39	52	9.1	41622	1	SCD25	AL118514 Streptomyces
C 40	52	9.1	143676	2	AP005182	Oryza sativa
C 41	51.8	9.0	2244	6	AR007590	Sequence
C 42	51.8	9.0	2334	6	AR007585	Sequence
C 43	51.8	9.0	2936	1	FSU90120	Flavobacterium
C 44	51.8	9.0	38332	1	SC9A4	Streptomyces
C 45	51.8	9.0	149256	2	AP003827	Oryza sativa

ALIGNMENTS

RESULT 1	
LOCUS	AP003416/c
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0456E05.
ACCESSION	AP003416
VERSION	AP003416.4
KEYWORDS	GI:20804922
SOURCE	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0456E05.
ORGANISM	Oryza sativa (japonica cultivar-group)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE	1

linear PLN 17-MAY-2002

CDS

```
complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/gene="P0456E05.7"
/notes="contains ESTs AU101665(R2414),D24709(R2594)
similar to Arabidopsis thaliana chromosome 5, At5g14420
unknown protein"
/codon_start=1
/protein_id="BAB92608.1"
/db_xref="GI:20804929"
/translation="MGGGSRSDRGSGHGRYGHSPFQQWQGGGGGGGGGGVGY
GDAHGGVYFAPQAGYAAPYATQPPMPAPPAQPPARAGGASKPLDRYS
RADYHGVQVDTALAQAGLESSNLIVGIDFTKSNWLGKSFHGRSLHLSNAPN
YEQAISIGQTLKDFEDNLIPCFGQDATTHQDVCEYDLRCNCFSEALSRYR
LPHLRLAGTSFAPINEMATIVQSGQYHVLIIADGQVRSVDYFASQGLSSQEQ
KSDAIVRASELPLSLVILVGVGDFGDMNKEFDVILIPSAFDFVFANFSEIMNKP
QSRKEAFAALMALPEIQYKATVGLGILGRSFKSPERVLPPPGGSHDAYSYGSK
FKPSYIPQSSSSSYPHYETAQSSPAVPSSTYDNOVPCILVNPDKMAFGCGHQ
NPKAFERLADRPVYWCIDRKASNPDLQDCGAGDVSSIETH"
51367..52887
/gene="P0456E05.8"
51367..52887
/gene="P0456E05.8"
/notes="hypothesized protein
similar to salt-inducible protein"
/codon_start=1
/protein_id="BAB92609.1"
/db_xref="GI:20804930"
/translation="NAASRLARKLPISLISKHORLISPEIDVEQAQSPASSIPLD
PSLPLAVSHLSPSPALPSPVHASTPAALRLRRARHPLRALDLHLALAA
SDSPAFRDLSTLSLASLAESRLPSRLLELVLPFCPCDDSTFACPELLTFE
RKALFAAGSDIPASDALASLRADSPLPAPFYNIILHALARLHEDTVRFYGE
MTSRHVPHDYVTINILNSCRAEGVDAAWFOEMRRWSCPTGYSFNTLMGFR
EGRCKGVAREMLQGLFGLSVASMEIWLNGRCGGEPKAAEVFEFLVDGVVPG
FDCLDLVESLCRVNVEKAVVELILERNVSLGLVPAGTVYLECLKGLDKACQ
MGRMVAEIVPDTISCNVFEALCEAGRTVDANRLQAKERQFADGFTYSIVOG
FGRCQIREGEAVLDEMDAGYVPIATYINRLDGLHMGSMQLOQKCSRHNAN"
Complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
Complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
```

Query Match

```
Best Local Similarity 30.4%; Score 174.6; DB 8; Length 173699;
Matches 257; Conservative 0; Mismatches 84; Indels 12; Gaps 2;

QY 76 CACACCAAGCGTCGGCACCACATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGG 135
DB 157591 CACTCCCCAACATAAAACAAATGGCTCCTACTACGAGGTGGACTACTGCTCGGAGGAGG 157532

QY 136 TGAGTTCGGTGGCCCC---CGGCCGGCTTCGCCGCCACGCGCGGGCGGTCCAGCAGCAGC 192
DB 157531 TGAGTTCGGTTCACCCCGCGGCTTCTTCGCGCGCGCGCGGGCGGTCCAGCAGCAGC 157472

QY 193 TCGTCAAGGAGAGTTCGAGGAGGTTCGACACGGTCTCACCGCGCGCGCGCGGCGGTCCAGCAGCAGC 252
DB 157471 TCGTCAAGGAGAGCTTCGAGGAGGTTCGACACGGTCTCACCGCGCGCGCGCGGCGGTCCAGCAGCAGC 157421

QY 253 ACCACCATGGTCACACCGCGCGCGCGGTTCGTTGTCGCGAGACACAGGTTCGAAGAGG 312
DB 157420 ACCACACACACACCGCGCGCGGTTCGTTGTCGCGAGACACAGGTTCGAAGAGG 157361

QY 313 ACATCAACACCTGTCACCGCGCGCGGTTCGTTGTCGCGAGACACAGGTTCGAAGAGG 372
DB 157360 ACTTCAACACCTGTCACCGCGCGGTTCGTTGTCGCGAGACACAGGTTCGAAGAGG 157301

QY 373 ACTGAGCGCGCGCGCGGTTCGTTGTCGCGAGACACAGGTTCGAAGAGG 425
DB 157300 ACTGATCGAACCTGCTGTGTGTACCGGTGTACGTACGTATATGTGTGCC 157248
```

RESULT 2
AC091071

LOCUS

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

DEFINITION

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

ACCESSION

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

VERSION

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

KEYWORDS

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

SOURCE

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

ORGANISM

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

REFERENCE

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

AUTHORS

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

TITLE

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

JOURNAL

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

REFERENCE

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

AUTHORS

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

TITLE

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

JOURNAL

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

COMMENT

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

BASE COUNT

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

ORIGIN

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Query Match

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Best Local Similarity

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Matches

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Conservative

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Mismatches

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Indels

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.

Gaps

AC091071 174218 bp DNA linear HTG 27-MAR-2001
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.


```
/gene="P0470A12.6"
Complement(join(34470..34734,34836..34921,35025..35117,
35201..35364,35440..35601,35682..35861,35978..36414,
36513..36568))
/gene="P0470A12.6"
/note="contains ESTs AU058030(C61369),AU088621(C61369)"
/codon_start=1
/product="putative RNA helicase"
/protein_id="BAB90282.1"
/db_xref="GI:20161357"
/translation="MNPYDLRYADPSSYRDRSLAGAPVLAASAPAAANPYAAAYAP
AAPAPAGGDFSRGGRGGGAGGGGGRGGGGGAGGGRGGGGGAGGGRGGGRDALS
LSLPKPFSLPFFKNEFYVECPAVQAMSDMDVSOYRRQRDITVEGHDPKPRFQF
ANFPDYCMQAIKASGFVEPTIQSQWPMALGRDMIGIAOTGSKTISYLLGPLVHQ
GAQPLEQDGPVILVLAPTRELAVQIOESGKFGSRYSTRSTCIYGGAPKGPORLD
RGVEIVATPCRLDMLGGHTNLRVTYVLVDEADRLMDGFEPOIRKIVAAQWLI
RDROTLYSASFAPREVESLARQFNQYKVIIGSPDLKANHSIQOITIEVISEHEKYP
RLSKLLSDMGSRILFIQTKKDKVTQLRQMDGMPALSIHGDKAQRDYVLAEP
KSGKSPMAATDVAARGLGMVTKLRIRL"
Join(40309..40472,40627..40822,40913..41002,41120..41176,
41290..41359,41646..41710,42307..42361,42603..42730,
42818..42906,43104..43173,43622..43723,43820..43909,
43983..44131,44519..44630,44713..44799,46231..46311,
46380..46604,46714..46836,46920..47036)
/gene="P0470A12.7"
Join(40309..40472,40627..40822,40913..41002,41120..41176,
41290..41359,41646..41710,42307..42361,42603..42730,
42818..42906,43104..43173,43622..43723,43820..43909,
43983..44131,44519..44630,44713..44799,46231..46311,
46380..46604,46714..46836,46920..47036)
/note="Contains ESTs
DI5893(C1477),C72138(E1066),AU101125(E1066)
similar to Arabidopsis thaliana chromosome 4, F7J7_90
unknown protein"
/codon_start=1
/protein_id="BAB90282.1"
/db_xref="GI:20161358"
/translation="MAAGGLPASATLLLVIAVAVAPLASVRPVSDAHRSAAREL
FAASPDGFLDTEYEAFTQILVEKDKGLIGKACFAAEKLAASSSPKADFLH
AARLSGKSDSVYDDVATRLKAVIKDTLSLELYYVGGLLSKKEGHNVLDPD
ADNTFHAIKALSDGWRDYDTNSAESSTFAAGTALSAVSLADSEYDSNIAVY
KNDVLFDTIKSYDDGTFYFDEKHDAEYKGPITTSASVVRGYSFAAASGLNI
FGKILGLAFVFLIGLPGSAKCFNIOVFVFLVLSLPSKFSLSLQKVEYTV
PGEAPFLNVLQVLSGDSKVIITETKELQFDLDNNHYLDIAPIDVYSLVFE
ISLQEQHETIAQVGTNTAEFTVGLIKVDAEIGISDNDAGTVEVKIDLOKDT
VLSANHLQLRUSFQSLPLGTFKPHOYFLKHKHDESKVHLFVVPVSGARQFVL
DFGLVKEFYLISGRYDLAVGDAAMENFLRALGHELDLPEAPEKAPKPAQAVD
PFSKFGPKFISHIFSRPKPKPELSFAFTGLTLPVIGFLGLRLGNLKNFSL
PAPAFASFLFHAGIGAVLLYLVFWIKDLFTTLKYLFLGVFLVFGHKLVSLSST
SAKOKTA"
Complement(join(47978..48061,48179..48239,48314..48378,
48478..48739,48837..48911,49020..49094,49242..49298,
49352..49625,49763..49964,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52059,52148..52561))
/gene="P0470A12.8"
Complement(join(47978..48061,48179..48239,48314..48378,
48478..48739,48837..48911,49020..49094,49242..49298,
49352..49625,49763..49964,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52059,52148..52561))

Query Match
Best Local Similarity 30.4%; Score 174.6; DB 8; Length 185095;
Matches 257; Conservative 0; Mismatches 84; Indels 12; Gaps 2;

76 CACACCAAGCGTCGGCACCACATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGAGG 135
|||||
27534 CACTCCCAACTAAAAACATGGCTCCTACAGGAGGTGGACTACTGCTCGGAGGAGG 27475
|||||

136 TGAGGTGCGTGGCC---CGGCGCGGCTTCGGCGCCACGGCGCGCTCCAGCAGCAGC 192
|||||
27474 TGAGGTGCGTGGCCCGCGGCGGCTTCTCGGCGCGGCTTCTCGGCGCGGCTCCAGCAGCAGC 27415
|||||
```

QY	76	CACACCAACGCTCGGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGAGGAGG	135
Db	156701	CACTCCCAACTAAACAAATGGCTCACTACAGAGAGTGGACTACTGCTCGAGGAGG	156760
QY	136	TGAGGTGCTGGCC---CGGCGGCTTCGGCCGCCACGGCGCGCTCCAGCAGCAGC	192
Db	156761	TGAGGTGCTGGCCACCGCGGCTTCCTCGCGCGCGCGCGCTCGAGCAGCAGC	156820
QY	193	TGCTCAAGGAGAGTTCGAGGAGGTCGACACGGTCTCAGCGCGCGCGCGCAACACCA	252
Db	156821	TGCTCAAGGAGAGTTCAGAGAGATCGACA-----GGTCCGGCTCCGCGCGCGCAC	156871
QY	253	ACCACATGCTCACACCGCGCGCGCTTCGCTGGCGGAGACCGAGGTCGAGAGG	312
Db	156872	ACCACAACACACCAACCGCAACGACTCTGATGCTGGCGGAGACCAAGTGGAGGAGG	156931
QY	313	ACATCAACACCTGCACCGGAGGTCACGAGCGAGGAGAGTTCCTCGCGCAGGGCTA	372
Db	156932	ACTTCAACACCTGCACCGGAGGTCGCGGAGGTCGCGGAGGAGAGTTCCTCGCTCAAGTCG	156991
QY	373	ACTGAGCGCGCGCGCGCGCATCCACCGCCGCTTCGCTGCTGCTGCTGCC	425
Db	156992	ACTGATCGAACCTGCTGTGTGTACCGGTGTACGTACGTATATGTGTGCTG	157044
RESULT 5			
LOCUS	SST242802	562 bp	mRNA linear PLN 27-NOV-1999
DEFINITION	Sporobolus stapfianus mRNA for putative glycine-rich protein.		
ACCESSION	AJ242802		
VERSION	AJ242802.1	GI:6478143	
KEYWORDS	glycine-rich protein.		
SOURCE	Sporobolus stapfianus.		
ORGANISM	Sporobolus stapfianus		
REFERENCE	Neale, A.D., Blomstedt, C.K., Bronson, P., Le, T.N., Guthridge, K., Evans, J., Gaff, D., and Hamill, J.D.		
AUTHORS	The isolation of lowly-transcribed genes which are induced during		
TITLE	dedication of the resurrection grass Sporobolus stapfianus		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 562)		
AUTHORS	Neale, A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-MAY-1999) Neale A.D., Biological Sciences, Monash		
LOCATION/Qualifiers	University, Wellington Rd Clayton, Victoria, 3168, AUSTRALIA		
FEATURES	1..562		
source	/organism="Sporobolus stapfianus"		
	/db_xref="taxon:56623"		
	/clone="SDG137c"		
	63..350		
CDS	/codon_start=1		
	/product="putative glycine-rich protein"		
	/protein_id="CAB61838.1"		
	/db_xref="GI:6478144"		
	/translation="MAHFKEVDYEVTSMKPGFGRHGGGVQFVYKETFEEVQV		
	TPRGSHGHGHRGNQHGSHGFQARETKFERHQHARRVPRAGQERPCQG"		
BASE COUNT	143 a	153 c	154 g 112 t
ORIGIN	1		
Query Match	25.1%; Score 143.8; DB 8; Length 562;		
Best Local Similarity	62.1%; Pred. No. 4e-12;		
Matches	321; Conservative	0; Mismatches	172; Indels 24; Gaps 5;
QY	79	ACCAGCGTCGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGGAGGAGTGA	138
Db	46	ACCAATCAAGATCAATCGCGCACTTCAAGAGATGTGCTACGAGAGGTGACGT	105
QY	139	GTGCTGTCGCCCGCGCGCTTCGCGCGCGCGGCTCCAGCGCGGCTCCAGCAGCAGCTGCTCA	198
Db	106	CGATGGCCAAACCGGGCTTCGGCGCCACCGCGCGCGCGCTCCAGCAGTTCGTCGTCA	165
QY	199	AGGAGAAGTTTCGAGGAGTGCACACGCTCTCACGCGCGCGCGCAACACCAACCA	253
Db	166	AGGAGAAGTTTCGAGGAGTGCAGAGGTACCCACCTGCGCGCGCGCGCAACCAACCA	225
QY	254	-----CCACCATGTTTCACACCGCGCGCGCTTCGCTGCTGCGCGAGACCGAGTTC	306
Db	226	GCCACCGCGCGCAACCAAGCCAGCGCTCCGCGCACTTCCAGCGCGCGCGCAACCA	285
QY	307	AAGAGGACATCAACACCTGCACCGCGCGAGTTCACGAGCGCGCGAGAGTTCCTCG	366
Db	286	A--GAGACATCAACACCGCGCGAGTTCACGAGCGCGCGAGTTCACGAGCGCGCG	343
QY	367	GGGCTAACTAGCGCGCGCGCGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTG	426
Db	344	GGGCTGACTGAGCAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	403
QY	427	TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	481
Db	404	TCATAAGTGTCTACCCGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT	463
QY	482	GTGACCGCAGCTCAGCTCTGCTACGAATTCACGACAAAGCTGCTGCTGCTGCTG	537
Db	464	ATGCACGCGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	522
QY	538	ACTTCTTCTGTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	574
Db	523	GCTTTTCGCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	559
RESULT 6			
LOCUS	AR135147	390 bp	DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 7 from patent US 6194550.		
ACCESSION	AR135147		
VERSION	AR135147.1	GI:14124052	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 390)		
AUTHORS	Gold, L., Tuerk, C., Pribnow, D. and Smith, J. Drew.		
TITLE	Systematic polypeptide evolution by reverse translation		
JOURNAL	Patent: US 6194550-A 7 27-FEB-2001;		
FEATURES	Location/Qualifiers		
source	1..390		
BASE COUNT	125 a	126 c	133 g 6 t
ORIGIN	1		
Query Match	10.9%; Score 62.6; DB 6; Length 390;		
Best Local Similarity	52.0%; Pred. No. 4.1;		
Matches	140; Conservative	0; Mismatches	129; Indels 0; Gaps 0;
QY	103	ACTACGAGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTGGGCGCGCGCGGCTTCG	162
Db	109	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	168
QY	163	GCGCGCACGCGCGCGCGCTCCAGCAGCTCTCAAGGAGAGTTCGAGGAGTTCGACA	222
Db	169	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	228
QY	223	CGGTCTCACGCGCGCGCGCAACCAACCAACCAACCAACCAACCAACCAACCAAC	282
Db	229	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	288
QY	283	TCGTGTCGCGGACGACGAGGTTCGAGAGGAGATCAACACCTGCACCGCGCGAGTCC	342
Db	289	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	346
QY	343	AGCGCAGGAGAGTTCCTTCGCGCGGCT 371	

```

Db      349  ACGACGACGACGACGACGACGCGCATGGTT 377

RESULT 7
SC3A7
LOCUS   Streptomyces coelicolor cosmid 3A7.          35100 bp      DNA       linear      BCT 12-MAY-2002
DEFINITION
ACCESSION   AL031155 AL645082
VERSION     AL031155.2 GI:20520754
KEYWORDS   3-oxoacid phosphate enol-lactone hydrolase/4-carboxymuconolactone
            decarboxylase; ABC transporter; beta-galactosidase; exonuclease;
            133 type transposase; secreted amidase; secreted chitinase;
            two-component sensor/regulator.
SOURCE     Streptomyces coelicolor A3(2).
ORGANISM   Streptomyces coelicolor A3(2).
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Streptomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE  1 (bases 1 to 35100)
AUTHORS   Redenbach, M., Kieser, H.M., Denapate, D., Eichner, A., Cullum, J.,
            Kinashi, H., and Hopwood, D.A.
TITLE      A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL    Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE    97000351
PUBMED     8843436
REFERENCE  2 (bases 1 to 35100)
AUTHORS   Oliver, K. and Harris, D.
TITLE      Unpublished
JOURNAL
PUBMED
REFERENCE  3 (bases 1 to 35100)
AUTHORS   Parkhill, J., Barrell, B.G. and Rajandream, M.A.
TITLE      Direct Submission
JOURNAL    Submitted (30-JUL-1998) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge,
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
            On May 9, 2002 this sequence version replaced gi:3367738.
COMMENT    Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC.
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
            numbered using the following system eg SC7B7.01c, SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous. The length
            in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon. Gene
            prediction is based on positional base preference in codons using a
            specially developed Hidden Markov Model (Krogh et al., Nucleic
            Acids Research, 22(22):4768-4778(1994)) and the Frameplot program
            of Bibb et al., Gene 30:157-66(1984) as implemented at
            http://www.nih.gov/jp/
            Jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
            correct initiation codon. Where possible we choose an initiation
            codon (atg, gtg, ttg or (att)) which is preceded by an upstream
            ribosome binding site sequence (optimally 5-13bp before the
            initiation codon). If this cannot be identified we choose the most
            upstream initiation codon.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone. It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions. Cosmid 3A7 lies
            between 10H5 and 4A2 on the AseI-B genomic restriction fragment.
            Location/Qualifiers
              1..35100
                /organism="Streptomyces coelicolor A3(2)"
                /strain="A3(2)"

misc_feature
RBS
gene
CDS

/db_xref="taxon:100226"
/clone="cosmid 3A7"
2..103
/note="overlap with cosmid 3A7 from 1 to 102"
complement(211..217)
/note="possible RBS upstream of SC3A7.01c"
complement(551..3229)
/gene="SC3A7.02c"
/note="SC06334"
complement(551..3229)
/gene="SC3A7.02c"
892 aa; similar to e.g. MALT-ECOLI Malt regulatory protein
(901 aa), fasta scores; opt: 540 z-score: 495.9 E();
2.3e-20, 26.2% identity in 936 aa overlap. contains
PS00017 ATP/GTP-binding site motif A (P-loop),
helix-turn-helix motif from aa 848 to 869 (Score 1239,
+3.37 SD) and P1am match to entry PF00196 GcrE, Bacterial
regulatory proteins, luxR family, score 50.40, E-value
3.8e-11"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="CAA20065.1"
/db_xref="GI:3367740"
/db_xref="SPTREMBL:O86603"
/translation="MGPSDGTGRSPADTGTDPDLPGLTREVLPSPATFLRRLI
VHLNQLRSPLTVNGACAGKTLTADNAGLRVPTAHLTVETGDHPGFWAIVFE
SLRVCCAGADTGPADPSGVQORLLASIAELNADSPVIVLDEYDMTNPIAE
OLEFVHHAGQGLRVLTTRTEPLPLHRYRSGELTEYRSELSFTTEAGALLRH
GLQPEYASALVDRTRGWAAGLRLSALAASTDPERYLKEFADHSTVADFLAEV
LKGRTDEEDLLRLSVDFCPDLADALTQRTDAGLLAGLHRENAFTERLHSMYR
LHPFREILRAHLRERLPGVEPELHRAARLRRSGFLPETLAHGAAGDFDLAAGL
VDLALQGLFTGLRSDALAQLFAGMPEARSASDLVRTAQALARDVNOGLRLRQA
RNHLSDGNSPGLAARLCALEALAEALTCGSPRAQAATAEELQEELPDHLDE
HPELSALLMTHVGSARLWREGREDDARAVLTRVAGPFGFTVLPVRSNGHLDDYL
TCWLGAERRALAAVSEARAGLPPRPGAGLRLVLAADVDRQELGRARVLLDYL
LPAREHDPVPTAGRALVRLLLDVGEDTDAAKAEPASTAVSPVPAEQAAVARS
ACLDAGPDEAVRLQSGDQPCVCAEAGIQVAAQRAAEELGAVRAGRAGPA
VSGAALVRARLAQESGDTVAACRLVARALLDRHLRRPFLQAGDWEPLATAPL
HDLAEGWLPGRPPAHGEPVPGADPAFLVPELSEGRERDRLRLRMKSTEEIANDL
YVSVINTVTKLSAYRKLAVNRGDVARRARDGLL"
complement(554..745)
/gene="SC3A7.02c"
/note="P1am match to entry PF00196 GcrE, Bacterial
regulatory proteins, luxR family, score 50.40, E-value
3.8e-11"
complement(3029..3052)
/gene="SC3A7.02c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
complement(3233..3236)
/note="possible RBS upstream of SC3A7.02c"
complement(3235..4149)
/gene="SC3A7.03c"
/note="SC06335"
complement(3235..4149)
/gene="SC3A7.03c"
/note="SC3A7.03c, questionable ORF, len: 304 aa; this ORF
is not predicted as coding by positional base preference"
/codon_start=1
/transl_table=11
/product="very hypothetical protein SC3A7.03c"
/protein_id="CAA20066.1"
/db_xref="GI:3367741"
/db_xref="SPTREMBL:O86604"
/translation="MPAEPVRIVRPRRPRFFGTCPKASWTAVMSDAVFALPGL
SSAATGRCRPRVGRPRRPPSARRSARWRRPPQSPQCPWATHPQGRQSPGSD
GLPRTAITSRRCPACQSRGRTRMRKRPACSRGMPVGDPEPAGSLGPNPVRPS
STLSLSCACAPSKRSPWAAEIAFLACPAGLEALPVRMPGRLPGAILVPAAGRRLPR
FPGSRQILRPTVAARSPPGGVTHRCWSACSGSVLSVPGARVGLRALTGLGPDVFFL
HGLAVKRLAVSKRRRFRPEADGPGSWA"
4175..4179
/note="possible RBS upstream of SC3A7.04"
RBS

```

gene	4186. .5340	/gene="SC3A7.04"	/note="SCO6336"	4186. .5340	/gene="SC3A7.04"	/note="SC3A7.04, questionable ORF, len: 384 aa; this ORF is not predicted as coding by positional base preference"	/codon_start=1	/transl_table=11	/product="hypothetical protein SC3A7.04"	/protein_id="CAA20067.1"	/db_xref="GI:3367742"	/db_xref="SPTREMBL:O86605"	/translation="MLFVIGWITAIHSRELOVPASAWTAAEYDGEIRDAANAWE ITGQVQVRLHNSQDGPARGRLPGRGQERGAERSQOQSRAPAEPSSTVPAQ GIPRVGHQPRRPVRRHNGHQVHLNLNLPVHGRPADPSHRPRLPDLIRVPRVPO DQRRGQPGPVLRRPDPHRLPHRITPDALNRTITGWAGSKTEANQLTAFY YDLLEAGQPTSTRYVGGKSGTAYTKITITAYDKLYRPTTELLPLDPSDIPVAGAPAT LOFEANLNGDTCRAPRNPSPASPRSATNPATPTRTTLGLGKAPYWTSCAYDDAG RRTQTATAAAGNTNTNRYKEGQPPALTELGLPCDGRHLRSCGGGQRS"	/complement(5244. .6176)	/gene="SC3A7.05c"	/note="SCO6337"	/complement(5244. .6176)	/gene="SC3A7.05c"	/note="SC3A7.05c, probable transposase, len: 310 aa; similar to transposase ORF's from many IS3 type insertion sequences e.g. Corynebacterium glutamicum TR:Q46069 (EMBL:X69104) IS3 related insertion element ORF2 (298 aa), fasta scores; opt: 610 z-score: 720.3 E(): 7.2e-33, 49.3% identity in 211 aa overlap"	/codon_start=1	/transl_table=11	/product="putative transposase"	/protein_id="CAA20068.1"	/db_xref="GI:3367743"	/db_xref="SPTREMBL:O86606"	/translation="MTFASGSCAGYSRSRGPASTGQWPARTRCPAGRGRTTSPSNAS FGTSRTRPTAGTSIAQVMTNKNVGTCLPVGNGFLYATVLDLICKRLKWTIPERS ATPVDLLQRTGAAQMPNKNVGTCLPVGNGFLYATVLDLICKRLKWTIPERS PTLVDLRAAAGAGDGLRGAISHGNGAAYVSKFEQVCSGELGVTGRTSRGAVGYS ADKAAESLNTMTKRGTLQGRKRWNGSARLAFVWRATYRTRRTTPASARSASPT NGDQLRPPHNDRCPSRQSPSSVRAGWNPFS"	/complement(6207. .6494)	/gene="SC3A7.06c"	/note="SCO6338"	/complement(6207. .6494)	/gene="SC3A7.06c"	/note="SC3A7.06c, probable transposase, len: 95 aa; similar to transposase ORF's from many IS3 type insertion sequences e.g. Corynebacterium glutamicum TR:Q46069 (EMBL:X69104) IS3 related insertion element ORF1 (97 aa), fasta scores; opt: 244 z-score: 383.2 E(): 4.3e-14, 46.2% identity in 93 aa overlap. Contains probable coiled-coil from 55 to 93 and helix-turn-helix motif from aa 25-46 (Score 1789, +5.28 SD)"	/codon_start=1	/transl_table=11	/product="putative transposase"	/query_match	10.7%; Score 61.4; DB 1; Length 35100;	Best Local Similarity	50.2%; Pred. No. 3.4;	Matches	152; Conservative	0; Mismatches	151; Indels	0; Gaps	0;	103	ACTACAGAGGTGGACTACTGCTCGAGAGAGGTGAGGTCGTGGCCCGCGCGCTCG 162	9769	ACTTCCTCTCTGGCAACGCGCGGAGATACCGGGCGGACGACGCGGAGGCATCG 9828	163	GC CGCGCAGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAGAAGTTCGAGAGGTGCGACA 222	9829	GCAGTGTCTGCCGACTCGGATCTCTGCTCGCAGCGCTGCGGGAGGGGGTTCAGCG 9888	223	CGGTCTACGCGCGCGCGCAACACCAACCAACCATGCTTACCACGCGCGCGCGCT 282
------	-------------	------------------	-----------------	-------------	------------------	--	----------------	------------------	--	--------------------------	-----------------------	----------------------------	---	--------------------------	-------------------	-----------------	--------------------------	-------------------	---	----------------	------------------	---------------------------------	--------------------------	-----------------------	----------------------------	---	--------------------------	-------------------	-----------------	--------------------------	-------------------	---	----------------	------------------	---------------------------------	--------------	--	-----------------------	-----------------------	---------	-------------------	---------------	-------------	---------	----	-----	--	------	--	-----	---	------	--	-----	--

```

strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous. The length
in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon. Gene
prediction is based on positional base preference in codons using a
specially developed Hidden Markov Model (Krogh et al., Nucleic
Acids Research, 22(22):4768-4778(1994)) and the Frameplot program
of Bibb et al., Gene 30:157-66(1984) as implemented at
http://www.nih.go.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gtg, ttg or (att)) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions. Cosmid StBAC28G1
overlaps cosmids StBAC20F6 and StBAC31E11.
Location/Qualifiers
1. .29080
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid BAC28G1"
2. .518
/note="previously sequenced DNA fragment. EMBL:X62373"
S.coelicolor genes for hydroxyacyl-CoA dehydrogenase"
2. .94
/note="nominal overlap with Streptomyces coelicolor cosmid
SCBAC20F6"
8. .73
/note="PS01162 Quinone oxidoreductase / zeta-crystallin
signature"
484. .5883
/note="previously sequenced DNA fragment. EMBL:X58833.1"
S.coelicolor 6 actva region genes of the actinorhodin
biosynthetic gene cluster"
502. .2103
/gene="SCBAC28G1.02"
/note="SCO5076; actVA1"
502. .2103
/gene="SCBAC28G1.02"
/note="SCBAC28G1.02"
/genes="SCBAC28G1.02, actVA1, probable integral membrane
protein, len: 533 aa; identical to previously sequenced
TR:Q53903 (EMBL:X58833) Streptomyces coelicolor 6 ActVA
region genes of the actinorhodin biosynthetic gene
cluster, ActVA1, 533 aa. Contains possible hydrophobic
membrane spanning regions"
/codon_start=1
/transl_table=11
/product="integral membrane protein"
/protein_id="CAC44189.1"
/db_xref="GI:14717083"
/db_xref="SPTREMBL:Q53903"
/translation="MTANPRGPGGDQGHPRWALIGVLVSLGILIDNTVLYNVL
RLTDEQGLGASHSOEVLNLTAFATLFTGWGLDGLRRRVLLGLGVLGUS
SLAGAYGSPQELIARACMGVSGSAVLPLSTLAAVFLRPRKALGRTWAAVSGFA
LIGIPVGTGLIAHFWMGSLVNVPLMAGLIVAVLPETRGTACGRVDAAGLLLS
IAGVPLVYAIIEAGRSGVTPRAVWAGLAGLGLVLTWHERRTPPEPSLELGFRRM
KAFSTAAVAGFVSFAMGFLTFSAFYLQSVRGYTLQAGCTVALAANVCGPLST
VLVRSIGPRNCAAGLAVTLCGVTFTQHPAVMLILVLEAAGVACVMPATAA
STWNAIPREKAGVASAMNVTROLGCAVGLSLMGAAVYRGIEDELAVLPSPARS
QAGESIDATLLAATRLGESLGVPAQRAQFADMLAAGAAVAVLGCALVLRWLPS
WTTPPPAGAVPGRHSDHLKVGQS"
2093. .2097
/genes="SCBAC28G1.02"
2105. .2500

```

```
(EMBL:X59833) Streptomyces coelicolor gene cluster, ActVA5, 381
of the actinorhizal biosynthetic gene cluster, ActVA5, 381
aa and similar to TR:Q9F0J3 (EMBL:AF218066) Streptomyces
araneae hydrolase NC_030405, 405 aa; fasta scores: opt: 1391
Z-score: 1521.5 bits: 290.3 E(): 3.9e-77; 56.522% identity
in 368 aa overlap"
/codon_start=1
/transl_table=11
/product="putative hydrolase"
/protein_id="CAC44193.1"
/db_xref="GI:14717087"
/db_xref="SPTREMBL:Q53907"
```

Query Match	10.1%;	Score 58.2;	DB 1;	Length 29080;
Best Local Similarity	49.5%;	Pred. No. 10;		
Matches 150;	Conservative 0;	Mismatches 153;	Indels 0;	Gaps 0;
109 AGGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCCCGCCGGCTTCGGCCGCC 168				
17986 AGGAGTTCGGAAGCTGGCCGGCGAGTTCGCGGCCGCTCCAGGACGGCGTCCGCCGACG 17927				
169 AGCGCGGGCGGCTCCAGCAGCACGTCGTTCAGGAGAAGTTCGAGGAGCTGCACACGGTCT 228				
17926 TCGTCGCCCGCTACCCGGGAGGCGCTCGAACGGCGCAACCGGTGTGCACGGCTGCACCG 17867				
229 CACGGCGCGCGCCCAACACCACCACCATGTTCCACCACGGCGCCACGGCTTCGTGG 288				
17866 ACCTCGGCGACCGTTCGCCCGCGCGACCGGGCGCCCCCGGCCAGTGTCCGCTGG 17807				
289 TCGCGGAGACCGAGGTTCGAAGAGGACATCAACACCTGCACCGCGGAGTTCACGAGCGCA 348				
17806 CCTTACCACATATGTGTGAGGAGACCGGCCCGCCACCGCGCCAGCCGACATCTCCGG 17744				
349 GGGAGAGCTTCCTTCGCGAGGGCTAACTAGCGCCCGCGGGCGGGGATCCAGCCCGTTC 408				
17746 AGCTGATCAGCGTGCACGCGGCCGTGTACGGGCTTCCCGGGCGGGCTCAGCGGGCATGC 17685				
409 GTG 411				
17686 GCG 17684				

RESULT 10	
LOCUS	SCC30
LOCUS	18420 bp DNA linear BCT 12-MAY-2002
DEFINITION	Streptomyces coelicolor cosmid C30.
ACCESSION	AL352972 AL645882
VERSION	AL352972.2 GI:20520831
KEYWORDS	ABC transporter ATP-binding subunit; ABC transporter integral membrane protein; arac-family regulatory protein; bifunctional protein (ribonuclease H/phosphoglycerate mutase); dehydrogenase; integral membrane protein; KMG/KDPG aldolase; transmembrane transport protein; transport protein; two-component response regulator; two-component sensor kinase.
SOURCE	Streptomyces coelicolor A3(2).
ORGANISM	Streptomyces coelicolor A3(2).
REFERENCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
AUTHORS	1 (bases 1 to 18420) Redenbach, M., Klesner, H. M., Denapaita, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D. A.
TITLE	A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL	Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE	97000351
PUBMED	8843436
REFERENCE	2 (bases 1 to 18420) Seeger, K. J. and Harris, D.
AUTHORS	Unpublished
JOURNAL	3 (bases 1 to 18420) Cerdeno, A. M., Parkhill, J., Barrell, B. G. and Rajandream, M. A.
AUTHORS	Direct Submission
TITLE	Submitted (14-APR-2000) Streptomyces coelicolor sequencing project.
JOURNAL	

Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: bartrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK
On May 9, 2002 this sequence version replaced gi:7619747.

COMMENT

Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7b7 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nh.gov.jp/jun/cgi-bin/frameplot.pl>.

CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid C30 Overlaps cosmid C75 on the AseI-C genomic restriction fragment.
Location/Qualifiers

FEATURES

source

1. 18420
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid C30"

complement(346..771)

/gene="rnpB"

complement(346..771)

/gene="rnpB"

/note="probable ribonuclease P RNA, (rnpB) (positions taken from Streptomyces lividans EMBL:SLXLANB)"

complement(681..1661)

/gene="SCC30.01c"

/note="SCC02293"

complement(681..1661)

/gene="SCC30.01c"

/note="SCC30.01c, possible integral membrane protein, len: 326 aa; similar to TR:BAA87714 (EMBL:AB016260)

Agrobacterium tumefaciens TIOF89 protein, 306 aa; fasta scores: opt: 396 z-score: 465.4 E(): 1.8e-18; 28.7%

identity in 286 aa overlap. Contains 2x Pfam match to entry PF00892 DUF6, integral membrane protein and possible hydrophobic membrane spanning regions"

/codon_start=1

/transl_table=11

/product="putative integral membrane protein"

/protein_id="CAB88171.1"

/db_xref="GI:7619748"

/db_xref="SPTREMBL:Q9L020"

/translation="MTARGWFLFLSMGVVWGVYPMIKVAVDEVSPGVVFCALGA ALLLPALRGDLGTVRRRKMPLAFVIEIGPWWTLDARHLSSSTAGLLIAGY PIVAVLLARFFGAERVGARRITGLGLAGVAVLTVPHLTGGDARSIAELVTVVGY ATAPLTAARHLKDVPTLQITLCLTAAVYVAPAFILRPATLPSPGEALAGLGLV CTAIAFVALELIKVGTPRAGVIYVNPVAVAGALLDEELTGLIGVAFITFLAG SVLTAAAGPGREARRVPMSTRQTSRAGGRVESLTGLPFRPGSTGQDG"

misc_feature

/gene="SCC30.01c"

/note="Pfam match to entry PF00892 DUF6, Integral membrane protein, score 51.40, E-value 1.9e-11"

misc_feature

/gene="SCC30.01c"

/note="Pfam match to entry PF00892 DUF6, Integral membrane protein, score 38.50, E-value 1.6e-07"

gene

/gene="SCC30.02c"

/note="SCC02294"

CDS

/gene="SCC30.02c"

/note="SCC30.02c, probable araC-family regulatory protein, len: 278 aa; similar to SW:ARAC_SALTY (EMBL:J01797) Salmonella typhimurium arabinose operon regulatory protein AraC, 281 aa; fasta scores: opt: 232 z-score: 270.4 E(): 1.3e-07; 26.7% identity in 270 aa overlap and to various Streptomyces coelicolor putative araC-family transcriptional regulators, e.g. TR:O50480 (EMBL:AL031031)

S. coelicolor putative transcriptional regulator SC7C7.17 (SC4H8.01), 288 aa; fasta scores: opt: 638 z-score: 605.1 E(): 3.3e-28; 43.8% identity in 292 aa overlap. Contains Pfam match to entry PF00165 HTH_AraC, Bacterial regulatory helix-turn-helix proteins, araC family and possible helix-turn-helix motif at residues 189..210 (+3.81 SD)"

/codon_start=1

/transl_table=11

/product="putative araC-family regulatory protein"

/protein_id="CAB88172.1"

/db_xref="GI:7619749"

/db_xref="SPTREMBL:Q9L019"

/translation="MAGAREQARNSRARGRGGPPDLTLTAREDRHYVAPHARDEVTV GVTGVEIVAYRGHHSVPGSIVVLEPGVGGPAAEGYSYALYAAARLLTGD TRTDVLPFRFVLDDEPLAALRAAHTDLARCDDPEAESRLPMLTAAARRSTA RAADTVPGAGRVARVVRDLADELLAPSLAALADGLSLRYLLRFRSTGHPY AWLAQHRVARARGLLDAGLRPAEVALYGFADQAHLTFRFRVLTVPATYRNSVQDR TG"

misc_feature

/gene="SCC30.02c"

/note="Pfam match to entry PF00165 HTH_AraC, Bacterial regulatory helix-turn-helix proteins, araC family, score 59.80, E-value 5.8e-14"

gene

/gene="SCC30.03"

/note="SCC02295"

2628..3389

/gene="SCC30.03"

2628..3389

/note="SCC30.03"

N-terminal region similar to TR:CAB52842 (EMBL:AL109848) Streptomyces coelicolor putative transcriptional regulator SC151.18, 548 aa; fasta scores: opt: 114 z-score: 134.1 E(): 5; 30.6% identity in 196 aa overlap. Contains possible coiled-coil region at approx residues 76..93"

/codon_start=1

/transl_table=11

/product="hypothetical protein SCC30.03"

/protein_id="CAB88173.1"

/db_xref="GI:7619750"

/db_xref="SPTREMBL:Q9L018"

/translation="MRIGELADTVGTVTRTVRHHQGLLPEPERLANGRYTTLRHV VLLAIRRLTELGLGLAEVRDLVADDAKDLAEVLELSDSLAROEAAIRERTTLRA LLGTGGVPGEGVSPLEAFAGIDVSGSPNALRDREMLVLESLVAPERRGLLA AMRGLGSPAAARARTTYELLDADAGADDPVAAEVARVLDACHPAELLPEGGDL DPGHLLRLYADFAPAQAQAEVRRMEIARRRS"

3386..4000

/gene="SCC30.04"

/note="SCC02296"

3386..4000

/gene="SCC30.04"

/note="SCC30.04, possible integral membrane protein, len: 204 aa; similar to TR:AAF12504 (EMBL:AE001863) Deinococcus radiodurans hypothetical 32.0 kD protein, 289 aa; fasta

scores: Opt: 178 z-score: 231.0 E(): 2e-05; 29.8% identity in 161 aa overlap. Contains possible hydrophobic membrane spanning regions. /codon_start=1 /transl_table=11 /product="possible integral membrane protein" /protein_id="CAB88174.1" /db_xref="GI:7619751" /db_xref="SPTREMBL:Q9L017" /translation="MRATGPAAPFYLLARHELRLWASLWLAARRTHDSAGGAAGFYARGGAMFGLAFVGVESVTMSVLLRDPAAHVLVDVYTVMVVGLHAASVVRP HYLDAAGTLRVALHVDVRIPLERTASVERELTHTPADGELDLAVGQSQTSTVVE LTGPVAHLSFFGRREVVRVRECHADDPGAFVAVELAREGHAVP" 4030..4034 RBS gene CDS		10.1%; Score 57.8; DB 1; Length 18420; Best Local Similarity 50.9%; Pred. No. 13; Matches 137; Conservative 0; Mismatches 132; Indels 0; Gaps 0; Query Match 129 GAGGAGTGAGTGGTGGCGCGCGGCTTCGGCGGCACGCGCGCGCGCTCCAGCAG 188 Db 6781 GCGGAGTGACCGTGGCGTCCGGTCCAGTCCCGCGCGCGCGCGCGCGCGCGCG 6840 Qy 189 CACGTGCTCAAGGAGAGTTCGAGGAGTTCGACACGGTCTCAGCGCGCGCGCGCGCGCG 248 Db 6841 GCCCGGTACTCGCGCCAGCTGTGTCGACGACGCGCGCGGAGTACTCGCGCGCGCTCCAC 6900 Qy 249 CACGACACCATGTCACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 308 Db 6901 GGCTTCCCGCGTCAACGCGGATCGAGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 6960 Qy 309 GAGGACATCAACACCTGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 368 Db 6961 CCGGACCGCGCGTGGCGCTTCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7020 Qy 369 GCTAACTGAGCG 397 Db 7021 CACGGACCG 7049	
RESULT 11 SCO18817/c LOCUS DEFINITION Streptomyces coelicolor A3(2) linear BCT 19-JUL-1999 actinorhodin biosynthetic genes, strain J1501. Y18817 ACCESSION Y18817.1 GI:5541917 VERSION methylintransferase; mini-circle protein; oxidoreductase; transcriptional regulator. KEYWORDS Streptomyces coelicolor A3(2). SOURCE Streptomyces coelicolor A3(2). ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces. 1 (bases 1 to 4851) Martinez-Costa, O.H., Martin-Triana, A.J., Martinez, E., Fernandez-Moreno, M.A. and Malpartida, F. An additional regulatory gene for actinorhodin production in Streptomyces lividans involves a LysR-type transcriptional regulator J. Bacteriol. 181 (14), 4353-4364 (1999) MEDLINE 99328982 PUBMED 10400594		2 (bases 1 to 4851) Malpartida, F. Direct Submission Submitted (01-MAR-1999) F. Malpartida, Centro Nacional de Biotecnologia, Microbiologia, Campus de la UAM 28049 Canto Blanco, Madrid, 28049, SPAIN FEATURES Location/Qualifiers source 1..4851 /organism="Streptomyces coelicolor A3(2)" /strain="J1501" /db_xref="taxon:100226" complement(32..529) /gene="ORF7" complement(32..529) /gene="ORF7" /codon_start=1 /transl_table=11 /product="putative mini-circle protein" /protein_id="CAB51132.1" /db_xref="GI:5541918" /translation="MPTPTTHRRDTPPPRTGGGEAETLGGFLDYLRSTIAGKVEDAP EPOVRTAAVASGTNLLGLNHLTHVERAVFLGEEVSWPATFRAASODGVADVAVIR EAVERNRVLDACDGLCAPLPRPGRPADSVRWALTHMVEETCRGHAGHAILREILID GATGR" complement(539..543) RBS /note="ORF7" complement(583..1209) /gene="ORF8" complement(583..1209) /gene="ORF8" /codon_start=1 /transl_table=11 /product="putative methyltransferase" /protein_id="CAB51133.1" /db_xref="GI:5541919" /translation="MSTTDVTFWDVGYAARPADAPRPNRVLYETVIGLPPGDALDL CGSGDALWLRGRWRTAVDISGVAARLGHARHGLDGLVADYRHDLRDSFPEG RFLVSAHYLTFPELDORASVLGCAHALRPGRLVVDHGSTAPWSWODDOPARHPA POEVAADIALDPAFWVERAEAPRTATGPGGRTAEVVDHVLVRRAD" complement(1215..1221) RBS /note="ORF8" 1344..1916 /gene="ORF8" 1344..1916 /gene="ORF9" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAB51134.1" /db_xref="GI:5541920" /translation="MDAATDDVLDVAVGPRLTRRERGITLAHLSAATGVSESTLSRL ESGRRATLELLPLARYDVPDLVGAPRTGDPRIHLKPVRFMGVYVPLSRPGG TOAFMIIPSRAPLETPOTHEGSEWLYVLGRLRLVGERDLTLSPGEAAEFDTSL PHWLGADGAVELLVPLGQVRAHVHG" complement(2168..3091) CDS /codon_start=1 /transl_table=11 /product="transcriptional regulator of LysR family" /protein_id="CAB51135.1" /db_xref="GI:5541921" /translation="MTDHHVDRDVEYFATVAEELHFTRAAELLYVSOPALSKQIRALE RQLGAELFHRGPRGVTLEAGTALLPHARMADWTEGAAVAEAAARRDTLVVGM TSPGRGLLPAIRSRFATAHPEAVLOLRMSWEDTAGLADTADYAVYVLPDADR YAVTVAEPRVLPDPSHPLAAREIDFADLGEFFLALPEAGVGLRDFWIALDERR ALDGPAAAPRVGAEISGTEYEALVAGLGLCATVGNAPLITLGGVVVTRVPRVGLSP SYALARREDEGRALVRGVEACRRVTDRA" 3180..3186 RBS /note="ORF11" 3192..4022 CDS /codon_start=1 /transl_table=11 /product="oxidoreductase of short-chain" /protein_id="CAB51136.1"	

[illegible]

Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
 clone:OJ1135_F06
 Published Only in Database (2001)
 2 (bases 1 to 102688)
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Direct Submission
 Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7458)
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1. 102688
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1135_F06"
 30773 a 22012 c 21866 g 27986 t 51 others

BASE COUNT
 ORIGIN

Query Match
 Best Local Similarity 9.88; Score 56.2; DB 2; Length 102688;
 Matches 124; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

Oy 172 GCGGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGTTCGACACGGTCTCAC 231
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 38485 GCGGCGCTCTCTGGACACGCCACCGGGGTGTCATCGCGGAGACCGCGTCTCGGCG 38426
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 232 GCGCGCGCGCCACACACACACACACCATGTTCACACGCGCGCCACGCGTTCGTGGTGC 291
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 38425 ACCGCGTGTCTATCTCTCCACAGTCACGCTGGCGGGACCGCGAGGCCGTCGCGACC 38366
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 292 GCGAGACACGAGGTCGGAAGAGGACATCAACACCTGCACCGCGCGAGTTCACAGACGAGGG 351
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 38365 GGCACCCCAAGTTCGCGACGGGGTCTCATCGGCGCGCGCGACCATCTCTCGGGAACG 38306
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 352 AGAGCTTCTCGCCAGGGTAACTAGACCGCGCGCGCGCATCTCACGCCCGTTC 408
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 38305 TCAGGATCGCGCGCGCGCAAGATCGGGCGCGGGTCTGCTGCTGCCACGTTGC 38249

RESULT 14'
 LOCUS
 AP004873
 DEFINITION
 Oryza sativa (japonica cultivar-group) DNA linear HTG 31-MAY-2002
 *** SEQUENCING IN PROGRESS ***; In ordered pieces.
 AP004873
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 HTG; HTGS_PHASE2.
 Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:P0453G09.
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1

REFERENCE
 AUTHORS
 TITLE
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
 clone:P0453G09

Search completed: June 16, 2003, 03:10:51
Job time : 1694.35 secs

AUTHORS Rafanan,E.R. Jr., Le,L., Zhao,L., Decker,H. and Shen,B.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAY-2000) Chemistry, University of California - Davis, One Shields Ave., Davis, CA 95616, USA

FEATURES
 source 1. 3164
 /organism="Streptomyces olivaceus"
 /strain="tu 2353"
 /db_xref="taxon:47716"
 339. .1958
 /gene="elmC"
 339. .1958
 /gene="elmC"
 /note="ElmC"
 /codon_start=1
 /transl_table=11
 /product="Tcm B2 oxygenase"
 /protein_id="AAF73050.1"
 /db_xref="GI:8118601"
 /translation="MDRIEIPVLVVGGLTGLAAAVFLRQGVDCILVERHRRSTFELT
 RASVINARTMELLNAGLEETVIDRSLHLEGRKRELGPADRIPMVVLRRARDLADI
 ERATVEPSLDADSVSTRAQWCGQDKLEILRDEAVRGADIRFHTRLDSFADQAD
 GYDAIVDRGTICARTAVRSYRLIAADGVRSIVRSLAGVTGTGHGSLGKAMSVLFQAD
 EPLVHGPRFVITYMANQAPGVLOTDFENRWIFGFCDAVGGDAAPFTGCADIVRT
 SLGILVLDVQLVQPMWEMSHVADSYSGRVFLAGDAAHVHPAGAFGANGGIQDAH
 NLWKLASVLHGSRASDALDITYHQERHPVGTETAEQAWTRITYRLDGDDELGRILVDI
 KVFAAGYRTYSVAGAAVPTAIPHEALTGLPGQRPVHWDHDKRKYSTVVDLAVDG
 FVLLARADGTPWADAARLAATTGIPLTAHVVGTLTDPADALAAATGLGEGAGALLL
 PDGFVARSDDTSADDPDAVLGDLARILART"
 1990. .2319
 /gene="elmH"
 1990. .2319
 /gene="elmH"
 /note="ElmH"
 /codon_start=1
 /transl_table=11
 /product="Tcm F1 monooxygenase"
 /protein_id="AAF73051.1"
 /db_xref="GI:8118602"
 /translation="MFRUSPDAAVTLNVNFTVSPDDQEKLYRLHLDVTERIHLPG
 FYSATFHLGNRRHVNVAQWSEETFRSMHALPELQEHFACRGIAIPLSVPCAVAL
 VFEEAAG"
 2316. .2636
 /gene="elmI"
 2316. .2636
 /gene="elmI"
 /note="ElmI"
 /codon_start=1
 /transl_table=11
 /product="Tcm F2 cyclase"
 /protein_id="AAF73052.1"
 /db_xref="GI:8118603"
 /translation="MTYRALMYLRMDPADAEVAAAFEDHDTTGLPPEELGTRRTTLFR
 FHOLYMHLEADDDVLPLKLEARDDPFRFKVNEEVGRYLTPVASDWSRLTDSKAEVIV
 TWEA"
 2648. .3079
 /gene="elmJ"
 2648. .3079
 /gene="elmJ"
 /note="may effect polyketide antibiotic elloramycin A
 biosynthesis"
 /codon_start=1
 /transl_table=11
 /product="ElmJ"
 /protein_id="AAF73053.1"
 /db_xref="GI:8118604"
 /translation="MTAOTPEPVIGVDVTPTRQGRTRALLTPTSAGATGGFLGTL
 DLGGERISHEVHPYSDKLYAVAGSVVVEVDGHEITLTGTDQALLTRGRMHRHFHRT
 EARPLVFIQISPLAPRDLPHGVTEPVPRQDAPPTVGGVR"
 487 a 1236 c 1003 g 437 t 1 others

BASE COUNT 9.5%; Score 54.6; DB 1; Length 3164;
 ORIGIN Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 05:14:00 ; Search time 1163.61 Seconds
(without alignments)
7989.092 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	417	72.6	417	13	BM501439
2	390	67.9	524	10	AW288876
3	387	67.4	500	10	AW787732
4	363	63.2	465	10	BE129897
5	340	59.2	523	10	AW331212
6	315	54.9	552	10	BE025302

7	283	49.3	436	10	AW787315	AW787315 945002E06
8	254	44.3	326	10	BE025303	BE025303 945028B09
9	251	43.7	443	10	BE129644	BE129644 945027E06
10	223	38.9	274	10	AW288875	AW288875 707009E07
11	210	36.6	476	10	BE519299	BE519299 945007B05
12	207	36.1	311	12	BF729420	BF729420 1000077C0
13	200	34.8	496	10	AW787314	AW787314 945002E06
14	178	31.0	648	14	BQ619167	BQ619167 RNOSQ04E0
15	178	31.0	648	14	BQ619315	BQ619315 RNOSQ06C0
16	178	31.0	648	14	BQ619337	BQ619337 RNOSQ06E0
17	178	31.0	648	14	BQ619383	BQ619383 RNOSQ07B0
18	178	31.0	648	14	BQ619390	BQ619390 RNOSQ07B0
19	174	30.3	553	11	AY104409	AY104409 Zea mays
20	167	29.1	433	10	BE225008	BE225008 945042F02
21	157	27.4	226	9	AI372183	AI372183 au02802.x
22	157	27.4	226	9	AI372183	AI372183 au02802.x
23	154	26.8	550	12	BG840383	BG840383 MEST12-H1
24	153	26.7	546	9	AA979839	AA979839 MEST2-B7
25	143	24.9	401	10	AW289056	AW289056 707005E07
26	139	24.2	466	9	AI964534	AI964534 496013D01
27	139	24.2	481	9	AI964458	AI964458 496013D01
28	139	24.2	539	9	AI855425	AI855425 603016F02
29	85	14.8	168	14	T12683	T12683 zEST00053-5
30	85	14.8	201	9	AA051890	AA051890 zEST000579
31	85	14.8	213	14	T18841	T18841 zEST00181-5
32	85	14.8	230	9	AA051874	AA051874 zEST00558
33	85	14.8	240	14	W21677	W21677 zEST00971 M
34	52	9.1	188	10	AW679347	AW679347 WSL_24_H0
35	52	9.1	287	10	BE125310	BE125310 DGL_19_C0
36	52	9.1	288	10	BE125500	BE125500 DGL_27_C0
37	52	9.1	348	10	BE593507	BE593507 WSL_100_B
38	52	9.1	399	13	BM318672	BM318672 P11_16_C0
39	52	9.1	466	10	BE597738	BE597738 P11_85_F0
40	52	9.1	485	10	BE599123	BE599123 P11_85_F0
41	52	9.1	509	14	BQ280894	BQ280894 WHE3006_B
42	52	9.1	513	14	BQ280709	BQ280709 WHE3004_A
43	52	9.1	516	10	BK364814	BK364814 P11_16_C0
44	52	9.1	519	10	AW680016	AW680016 WSL_34_H1
45	52	9.1	522	10	AW679969	AW679969 WSL_34_B0

ALIGNMENTS

RESULT 1
BM501439 417 bp mRNA linear EST 14-FEB-2002
LOCUS PAC000000000593 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
DEFINITION BM501439
ACCESSION BM501439
VERSION BM501439.1 GI:18661517
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 417)
Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and Jung,R.

AUTHORS

Maize opaque endosperm mutations create extensive changes in patterns of gene expression
Unpublished (2002)
Contact: Jung R

TITLE

Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA

JOURNAL

Tel: 515 270 5934
Fax: 515 254 2619

COMMENT

Email: rudolf.jung@pioneer.com.
Location/Qualifiers
1..417
/organism="Zea mays"
/db_xref="taxon:4577"

FEATURES

source

Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. .500

FEATURES

source

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
Note: "Organ: tassel, kernal, silk, husk, root, leaf; Vector: pGAD10; Site: 1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."
BASE COUNT 97 a 159 c 156 g 88 t

Query Match 67.4%; Score 387; DB 10; Length 500;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 62 GCGCTGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTGGACTA 121
DB 1 GCGCTGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTGGACTA 60
QY 122 CTGCTCGGAGAGGTGAGTGGTGGCCCGCGCGGTTCGGCCGCCACGCGCGCGCGT 181
DB 61 CTGCTCGGAGAGGTGAGTGGTGGCCCGCGCGGTTCGGCCGCCACGCGCGCGCGT 120
QY 182 CCAGCAGCAGCTGCTCAAGGAGAAGTTCCGAGGAGGTGCGACACGGTCTCACGCGCGCGC 241
DB 121 CCAGCAGCAGCTGCTCAAGGAGAAGTTCCGAGGAGGTGCGACACGGTCTCACGCGCGCGC 180
QY 242 CAACACACACACCAACCATGTCACACGCGCGCGCGGTTCGTGTGTCGCGCAGACGAC 301
DB 181 CAACACACACCAACCATGTCACACGCGCGCGCGGTTCGTGTGTCGCGCAGACGAC 240
QY 302 GGTGCAAGAGACATCAACCTTCACCGCGCGCGGTTCGAGGAGGTCCAGAGCGAGGAGCTTCCT 361
DB 241 GGTGCAAGAGACATCAACCTTCACCGCGCGCGGTTCGAGGAGGTCCAGAGCGAGGAGCTTCCT 300
QY 362 CGCCAGGCGTAACTGAGCGCGCGCGCGCGGTTCGAGGAGGTCCAGAGCGAGGAGCTTCCT 421
DB 301 CGCCAGGCGTAACTGAGCGCGCGCGCGCGGTTCGAGGAGGTCCAGAGCGAGGAGCTTCCT 360
QY 422 TGCCTTATGATGCTGTGTTGACTGTTGTTTTCAGGGTTCATCGTACTTGGCTATCGTAC 481
DB 361 TGCCTTATGATGCTGTGTTGACTGTTGTTTTCAGGGTTCATCGTACTTGGCTATCGTAC 420
QY 482 GTGCACGCACTAGCTCTCTACGAAATACGACAAATAGCTCGTACCTGAATAAACTT 541
DB 421 GTGCACGCACTAGCTCTCTACGAAATACGACAAATAGCTCGTACCTGAATAAACTT 480
QY 542 CTTGCGTAAT 550
DB 481 CTTGCGTAAT 489

RESULT 4

BE129897
LOCUS 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897
VERSION BE129897.1 GI:8577260
KEYWORDS Zea mays.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 465)
Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1999)

Contact: Walbot V
Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221

Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.

Location/Qualifiers
1. .465

FEATURES

source

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
Note: "Organ: tassel, kernal, silk, husk, root, leaf; Vector: pGAD10; Site: 1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."
BASE COUNT 88 a 151 c 149 g 77 t

Query Match 63.2%; Score 363; DB 10; Length 465;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 57 CACGGCGCTGGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTTACCAGGAGGTG 116
DB 1 CACGGCGCTGGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTTACCAGGAGGTG 60
QY 117 GACTACTGCTCGGAGAGGTGAGTGGTGGCCCGCGCGGTTCGGCGCGCACGCGCGC 176
DB 61 GACTACTGCTCGGAGAGGTGAGTGGTGGCCCGCGCGGTTCGGCGCGCACGCGCGC 120
QY 177 GCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGGTGCGACACGGTCTCACGCGCC 236
DB 121 GCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGGTGCGACACGGTCTCACGCGCC 180
QY 237 GCGCGCAACCAACCAACCAACCATGCTCACCACGCGCGCGGTTCGTGTGTCGCGCAG 296
DB 181 GCGCGCAACCAACCAACCAACCATGCTCACCACGCGCGCGGTTCGTGTGTCGCGCAG 240
QY 297 ACCAGGGTCGAGAGGACATCAACACTGCAACCGCGCGAGGTCCAGAGCGCAGGAGAGC 356
DB 241 ACCAGGGTCGAGAGGACATCAACACTGCAACCGCGCGAGGTTCACAGCGCAGGAGAGC 300
QY 357 TTCCTCGCGCAGGGCTAACTGAGCGCGCGCGCGGTTCACCGCGCGGTTCGTCGCTTGC 416
DB 301 TTCCTCGCGCAGGGCTAACTGAGCGCGCGCGCGGTTCACCGCGCGGTTCGTCGCTTGC 360
QY 417 CTGGTGGCTTATGATGCTGTGTTGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 476
DB 361 CTGGTGGCTTATGATGCTGTGTTGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 420
QY 477 CGTACGTGCGCAGCACTCAGCTCCTGTACGAAATACGACAAATAGC 521
DB 421 CGTACGTGCGCAGCACTCAGCTCCTGTACGAAATACGACAAATAGC 465

RESULT 5

D	b	362	GCGTTAAGTATCTATGCTTGACTGGTTGTGCAGGGTCATCTGTA	CTTGGCTATCGTAGC 421
Q	y	483	TGCAGCACTCAGCTCCTCTCTACGAATTAGACAATAAGCTCGTGACCTGAATAAACTTC	542
D	b	422	TGCAGCACTCAGCTCCTCTAGCAATTAGACAATAAGCTCGTGACCTGAATAAACTTC	481
Q	y	543	TTTCGTAATACTAA	555
D	b	482	TTTCGTAATACTAA	494
RESULT 6				
BE025302				
LOCUS				
DEFINITION				
552 bp mRNA linear EST 07-JUN-2000				
(SK) Zea mays cDNA, mRNA sequence.				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Zea mays.				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC				
clade; Panicoideae; Andropogoneae; Zea.				
1 (bases 1 to 552)				
Walbot.V.				
Maize ESTs from various cDNA libraries sequenced at Stanford				
University				
Unpublished (1999)				
Contact: Walbot V				
Department of Biological Sciences				
Stanford University				
855 California Ave, Palo Alto, CA 94304, USA				
Tel: 650 723 2237				
Fax: 650 725 8221				
Email: walbot@stanford.edu				
Plate: 945028 row: B column: 09.				
Location/Qualifiers				
1..552				
/organism="Zea mays"				
/cultivar="W23"				
/db_xref="taxon:4577"				
/clone_lib="945 - Mixed adult tissues from Walbot lab,				
same as 707 (SK)"				
/tissue_type="tassel, kernal, silk, husk, root, leaf"				
/dev_stage="fully-grown"				
/lab_host="DH10B"				
/note="Organ:tassel, kernal, silk, husk, root, leaf;				
Vector:pGDH10; site.1:EcoRI; CDNA library from fully				
differentiated maize tissues from an active Mutator plant				
Tissue ratio is 4:2::1:1:1 (tassel, kernel, silk, husk,				
root, leaf). Unidirectionally cloned. New library number				
given to library 707 for additional sequencing."				
BASE COUNT				
ORIGIN				
Query Match 54.9%; Score 315; DB 10; Length 552;				
Best Local Similarity 99.2%; Fred. No. 0;				
Matches 515; Conservative 0; Mismatches 4; Indels 0; Gaps 0				
Q	y	49	CGAGCTCACAGGGCGCTGCCACACAGACACCAAGCGTCGCACCAATGGCTTACTACC	108
D	b	1	CGAGCTCACAGGGCGCTGCCACACAGACACCAAGCGTCGCACCAATGGCTTACTACC	60
Q	y	109	AGGAGGTGACTACTGCTCGGAGGAGGTAGCTGGCCGCCGCGCTTCGGCGCGC	168
D	b	61	AGGAGGTGACTACTGCTCGGAGGAGGTAGCTGGCCGCCGCGCTTCGGCGCGC	120
Q	y	169	ACGGCGGGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAGGTCGACACGGTCT	228
D	b	121	ACGGCGGGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAGGTCATACGGTCT	180
Q	y	229	CACGCGCGGCGGCCAACACACACCAACCATGGTCCACACGGCGCCACGGCTTCGTGG	288

AW331212	523 bp	mRNA	linear	EST 31-JAN-2000
LOCUS	707049E04.x1 707 - Mixed adult tissues from Walbot lab (SK)	Zea mays cDNA, mRNA sequence.		
DEFINITION	AW331212			
ACCESSION	AW331212.1	GI:6827569		
VERSION	AW331212.1			
KEYWORDS	EST.			
SOURCE	Zea mays.			
ORGANISM	Zea mays			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 523)			
TITLE	Walbot.V.			
JOURNAL	Maize ESTs from various cDNA libraries sequenced at Stanford University			
COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855, California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 707049 row: E column: 04. Location/Qualifiers 1..523 /organism="Zea mays" /cultivar="W23" /db_xref="taxon:4577" /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)" /tissue_type="tassel, kernel, silk, husk, root, leaf" /dev_stage="adult" /lab_host="DH10B"			
FEATURES				
source				
BASE COUNT	125 a 158 c 149 g 91 t			
ORIGIN				
Query Match	59.2%; Score 340;	DB 10;	Length 523;	
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches	490;	Conservative 0;	Mismatches 3;	Indels 0;
Gaps	0;			
Qy	63	C G C T G C G C A C A C A C A C A C G G T C G C G C A C C A A T G C T T A C C A G A G G T G G A C T A C	122	
Dd	2	C G T G C G C A C A C A C A C A C G T C G C C A C C A T G C T T A C C A G A G G T G G A C T A C	61	
Qy	123	T G C T C G G A G A G G T G A G G T C G G T G C C C C G C C G G C C C A C G G C C C A C G G C G C G C G C	182	
Dd	62	T G C T C G G A G A G G T G A G G T C G G T G C C C C G C C G C G C T T C G C G C C C A C G C G C G C G C G C	121	
Qy	183	C A C A C A C G T C G T C A A G S A A G T T C G A G A G G T C G A C A C G G T C T A C G C G C C G C G C C	242	
Dd	122	C A G C A C A G T G T C A A G A G A A G T T C A G A G A G T T C A C A C G G T C T A C G C C C G C G C C	181	
Qy	243	A A C C A C C A C C A C C A C C A T G G T C A C A C G C G C C C A C G G C T T C G T G T G C C G A G A C C A G G	302	
Dd	182	A A C C A C C A C C A C C A C C A T G G T C A C A C G C G C C C A C G G C T T C G T G T G C C G A G A C C A G G	241	
Qy	303	G T C G A G A G A C A T C A A C A C T G C A C C G C G A G G T T C C A C A G A C G C A G G A G A G C T T C C T C	362	
Dd	242	G T C G A G A G A C A T C A A C A C T G C A C C G C G A G T T C C A C A G A C G C A G G A G A G C T T C C T C	301	
Qy	363	G C C A G G G C T A A C T G A C C G C C C G G C C G G A T C C A C G C C G T T C G T G C T T G C C T G C G G T	422	
Dd	302	G C C A G G G C T A A C T G A C C C C C G G G C C G G C A T C C A C G C C G T T C G T G C T T G C C T G C G G T	361	
Qy	423	G C C T A G T A T G T C T G T G T T G A C T T G G T T G T T C A G G G T C A T C G T A C T T G G C T A T C G T A C G	482	

root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

BASE COUNT 59 a 112 c 112 g 43 t
ORIGIN

Query Match 44.3%; Score 254; DB 10; Length 326;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 116 GGACTACTCTCGGAGGAGTGGAGTCTCGTGGCCCGCGGCTTCGGCCGCCACCGCGG 175
|||||
DB 22 GGACTACTCTCGGAGGAGTGGAGTCTCGTGGCCCGCGGCTTCGGCCGCCACCGCGG 81
|||||

QY 176 CGGCGTCCAGCAGCAGTCTCAAGAGAGAGTTCGAGGAGTTCGACACAGTCTCAGCGC 235
|||||
DB 82 CGGCGTCCAGCAGCAGTCTCAAGAGAGAGTTCGAGGAGTTCGACACAGTCTCAGCGC 141
|||||

QY 236 CGGCGCCACACACACACACATGTTCCACACGGCGCCACGGTTCTGTTGGTGGCGGA 295
|||||
DB 142 CGGCGCCACACACACACACATGTTCCACACGGCGCCACGGTTCTGTTGGTGGCGGA 201
|||||

QY 296 GACGAGGTCGAGAGGAGATCAACACCTGCACCGCGGAGTCCAGCAGCGCAGGAGAG 355
|||||
DB 202 GACGAGGTCGAGAGGAGATCAACACCTGCACCGCGGAGTTCACGAGGTCGAGGAGAG 261
|||||

QY 356 CTTCCTCGCAGGAGTCACTGAGCCCGCGCGCGCGCATCCACGCCGTTCTGTTGGT 415
|||||
DB 262 CTTCCTCGCAGGAGTCACTGAGCCCGCGCGCGCGCATCCACGCCGTTCTGTTGGT 321
|||||

QY 416 CTTGC 420
|||||
DB 322 CTTGC 326

RESULT 9
BE129644
LOCUS
DEFINITION 945027E06.X2 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.

ACCESSION BE129644
VERSION BE129644.1 GI:8577007
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 443)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945027 row: E column: 06.

FEATURES
source
1. .443
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site_1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant."

Tissue ratio is 4:2:1:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

BASE COUNT 78 a 144 c 148 g 73 t
ORIGIN

Query Match 43.7%; Score 251; DB 10; Length 443;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 TTACTACAGAGGTGGACTACTCTCGGAGGAGTGGAGTCTCGTGGCCCGCGGCTTC 160
|||||
DB 43 TTACTACAGAGGTGGACTACTCTCGGAGGAGTGGAGTCTCGTGGCCCGCGGCTTC 102
|||||

QY 161 CGGCGCCACACACACACACATGTTCCAGCAGCAGTCTCAAGAGAGTTCGAGGAGTCCA 220
|||||
DB 103 CGGCGCCACACACACACACATGTTCCAGCAGCAGTCTCAAGAGAGTTCGAGGAGTCCA 162
|||||

QY 221 CAGGTTCTACCGCGCGCGCGCGCAACACACACACATGTTCCAGCAGCGCGCCACCG 280
|||||
DB 163 CAGGTTCTACCGCGCGCGCGCGCAACACACACACATGTTCCAGCAGCGCGCCACCG 222
|||||

QY 281 CTTCTGTGTGCGGAGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGAGTCCA 340
|||||
DB 223 CTTCTGTGTGCGGAGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGAGTCCA 282
|||||

QY 341 CGAGCGCAGGAGAGGTTCTCTCGCAGGAGTCACTGAGCGCGCGCGCGCGCATCCAC 400
|||||
DB 283 CGAGCGCAGGAGAGGTTCTCTCGCAGGAGTCACTGAGCGCGCGCGCGCGCATCCAC 342
|||||

QY 401 GCCGTTCTGCTGCTGCCTGCCTTATGTATGCTGTGCTGCTGCTGCTGCTGCTG 453
|||||
DB 343 GCCGTTCTGCTGCTGCCTGCCTTATGTATGCTGTGCTGCTGCTGCTGCTGCTG 395
|||||

RESULT 10
AW288875
LOCUS
DEFINITION 707009E07.x3 707 - Mixed adult tissues from Walbot lab (SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW288875
VERSION AW288875.1 GI:6695662
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 274)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES
source
1. .274
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site_1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant."

differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned.

BASE COUNT 57 a 91 c 93 g 33 t

Query Match 38.9%; Score 223; DB 10; Length 274;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CAAGCATTTCAGCTTCACACGGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 99

DB 1 CAAGCATTTCAGCTTCACACGGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 60

QY 100 CTTACTACAGGAGTGGACTACTCTCGGAGGAGTGGCTCGGCGCGCGCGCT 159

DB 61 CTTACTACAGGAGTGGACTACTCTCGGAGGAGTGGCTCGGCGCGCGCGCT 120

QY 160 TCGGCGCGCACGGCGCGCTCGGCACACACACACAAAGCTTCGAGGAGTGG 219

DB 121 TCGGCGCGCACGGCGCGCTCGGCACACACACACAAAGCTTCGAGGAGTGG 180

QY 220 ACACGGTCTCAGCGCGCGCGCACACACACACCAATGG 262

DB 181 ACACGGTCTCAGCGCGCGCGCACACACACCAATGG 223

RESULT 11

BE519299

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BE519299 945007B05.x1 945 - Mixed adult tissues from Walbot lab, same as 707

(SK) Zea mays cDNA, mRNA sequence.

BE519299 BE519299.1 GI:9743151

EST.

Zea mays.

Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 476)

Walbot.V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 945007 row: B column: 05.

Location/Qualifiers

1. 476

/organism="Zea mays"

/cultivar="W23"

/db_xref="taxon:4577"

/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"

/tissue_type="tassel, kernel, silk, husk, root, leaf"

/dev_stage="fully-grown"

/lab_host="DH10B"

/note="Organ: tassel, kernel, silk, husk, root, leaf; vector: pGAD10; Site: 1; EORI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

99 a 145 c 146 g 86 t

Query Match 36.6%; Score 210; DB 10; Length 476;

BASE COUNT

ORIGIN

Best Local Similarity 98.9%; Pred. No. 0;

Matches 460; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 GCCTTACTACAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCGCGCGG 157

DB 12 GCCTTACTACAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCGCGCGG 71

QY 158 CTTGCGCGCACGGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 217

DB 72 CTTGCGCGCACGGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 131

QY 218 CGACACGGTCTCAGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 277

DB 132 CGACACGGTCTCAGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 191

QY 278 CGGCTTGTGCTGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 337

DB 192 CGGCTTGTGCTGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 251

QY 338 CCAGCAGCGGAGGAGGTTCCTGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 397

DB 252 CCAGCAGCGGAGGAGGTTCCTGCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 311

QY 398 CAGCGCGGTTCGCTTGTGCTGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 457

DB 312 CAGCGCGGTTCGCTTGTGCTGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 371

QY 458 GGTGATGCTTGTGCTTGTGCTGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 517

DB 372 GGTGATGCTTGTGCTTGTGCTGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAGGT 431

QY 518 AAGCTGCTGACCTGATAAATCTCTGTAATCTTAAAAA 562

DB 432 AAGCTGCTGACCTGATAAATCTCTGTAATCTTAAAAA 476

RESULT 12

BF729420

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BF729420 1000077C04.x2 1000 - Unigene I from Maize Genome Project Zea mays

cDNA, mRNA sequence.

BF729420 BF729420.1 GI:12047281

EST.

Zea mays.

Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 311)

Walbot.V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1000077 row: C column: 04.

Location/Qualifiers

1. 311

/organism="Zea mays"

/db_xref="dbEST:945032C12.Y1"

/db_xref="taxon:4577"

/clone_lib="1000 - Unigene I from Maize Genome Project"

/note="This library represents the unique ESTs found in the first round of EST sequencing at Stanford University for the maize genome project. Sequences are present from libraries 486, 487, 496, 603, 606, 614, 618, 660, 683, 687, 707, and 945. Contigs were assembled using TIGR's

CAP program and a representative EST from each contig was selected for the Unigene set. All singlets were also selected."

```

BASE COUNT      53 a 105 c 107 g 46 t
ORIGIN

Query Match      36.1%; Score 207; DB 12; Length 311;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 183 CAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGTTCGACACGGTCTCAGCGCCGCCGCC 242
      |||||
DB 54 CAGCAGCAGCTGCTCAAGGAGAAGTTTCGAGGAGTTCGACACGGTCTCAGCGCCGCCGCC 113
      |||||
QY 243 AACCCACACACCATGCTCACCACGCGCGCCGCTTCGTCGCGGAGACACAGG 302
      |||||
DB 114 AACCCACACACCATGCTCACCACGCGCGCCGCTTCGTCGCGGAGACACAGG 173
      |||||
QY 303 GTCGAAGAGACATCAACACCTGCACCGCGGAGTTCACGAGCGCAGGAGAGCTTCCTC 362
      |||||
DB 174 GTCGAAGAGACATCAACACCTGCACCGCGGAGTTCACGAGCGCAGGAGAGCTTCCTC 233
      |||||
QY 363 GCCAGGCTAACTAGACGCGCGCGCGCGCGCATCCACGCGCTTCGTCGCTGCGCT 422
      |||||
DB 234 GCCAGGCTAACTAGACGCGCGCGCGCGCGCATCCACGCGCTTCGTCGCTGCGCT 293
      |||||
QY 423 GCCTATGATGCTGTG 440
      |||||
DB 294 GCCTATGATGCTGTG 311
      |||||

```

```

RESULT 13
AW787314.
LOCUS      496 bp mRNA linear EST 16-MAY-2000
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787314
VERSION    AW787314.1 GI:7844111
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
REFERENCE  1 (bases 1 to 496)
AUTHORS    Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 945002 row: E column: 06.
FEATURES   source
            1..496
            Location/Qualifiers
            1..496
            /organism="Zea mays"
            /cultivar="W23"
            /db_xref="taxon:4577"
            /clone_lib="945 - Mixed adult tissues from Walbot lab,
            same as 707 (SK)"
            /tissue_type="tassel, kernal, silk, husk, root, leaf"
            /dev_stage="fully-grown"
            /lab_host="DH10B"
            /note="Organ: tassel, kernal, silk, husk, root, leaf;
            Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
            differentiated maize tissues from an active Mutator plant.
            Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
            root, leaf). Unidirectionally cloned. New library number
            given to library 707 for additional sequencing."

```

```

BASE COUNT      97 a 156 c 153 g 90 t
ORIGIN

Query Match      34.8%; Score 200; DB 10; Length 496;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 249 CACCACCACTGTTTCACACGCGCGCCGCTTCGTCGCGGAGACAGGTCGAA 308
      |||||
DB 183 CACCACCACTGTTTCACACGCGCGCCGCTTCGTCGCGGAGACAGGTCGAA 242
      |||||
QY 309 GAGGACATCAACCTGTCACCGCGGAGTTCACGAGCGCAGGAGAGCTTCCTGCCACG 368
      |||||
DB 243 GAGGACATCAACCTGTCACCGCGGAGTTCACGAGCGCAGGAGAGCTTCCTGCCACG 302
      |||||
QY 369 GCTAACTGAGCGCGCGCGCGCATCCACGCGCTTCGTCGCTGCGTGCCCTTA 428
      |||||
DB 303 GCTAACTGAGCGCGCGCGCGCATCCACGCGCTTCGTCGCTGCGTGCCCTTA 362
      |||||
QY 429 TGTATGCTGCTGTTGACTGTTGTTTCAGGTCATCGTACTTGGCTATCGTACGTCACG 488
      |||||
DB 363 TGTATGCTGCTGTTGACTGTTGTTTCAGGTCATCGTACTTGGCTATCGTACGTCACG 422
      |||||
QY 489 CACTCAGCTCCTGTACGAAATACGACATTAAGCTCGTACCTGAATAAACTTCTTCGTA 548
      |||||
DB 423 CACTCAGCTCCTGTACGAAATACGACATTAAGCTCGTACCTGAATAAACTTCTTCGTA 482
      |||||
QY 549 AT 550
      ||
DB 483 AT 484
      ||

```

```

RESULT 14
BQ619167.
LOCUS      648 bp mRNA linear EST 27-JUN-2002
DEFINITION RNOSEQ4E05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
mays cDNA clone RNOSEQ4E05_SK.ab1 similar to No homology, mRNA
sequence.
ACCESSION BQ619167
VERSION    BQ619167.1 GI:21621161
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
REFERENCE  1 (bases 1 to 648)
AUTHORS    Wang,H. and Bohnert,H.J.
TITLE      Genomics of plant stress tolerance
JOURNAL    Unpublished (2002)
COMMENT    Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnerlab@life.uiuc.edu.
FEATURES   source
            1..648
            Location/Qualifiers
            1..648
            /organism="Zea mays"
            /db_xref="taxon:4577"
            /clone_lib="RNOSEQ4E05_SK.ab1"
            /tissue_type="Salt stressed Zea mays roots cDNA library"
            /dev_stage="2 weeks old"
            /note="Vector: pBluescript SK+; Stressed 24 hours at 150
            mM NaCl"

```

```

BASE COUNT      172 a 193 c 183 g 100 t
ORIGIN

Query Match      31.0%; Score 178; DB 14; Length 648;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 328; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

[illegible]

Search completed: June 16, 2003, 08:38:46
Job time : 1170.61 secs

[illegible]

	Query Match	31.0%	Score 178;	DB 14;	Length 648;
	Best Local Similarity	99.1%;	Pred. No. 0;		
	Matches 328;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	244	ACACACCACACCATGTTCCACCGGGCGGCACGGCTTCGTGTGGCGAGACACGAGG	303		
dbb	233	ACACACCACACCATGTTCCACCGGGCGGCACGGCTTCGTGTGGCGAGACACGAGG	292		
QY	304	TCAAGAGGACATCAACACTCCACCGGGCGAGGTCACGAGCGGACGAGAGCTTCCTCG	363		
dbb	293	TCGANGGACATCAACACTCCACCGGGCGAGGTCACGAGCGGACGAGAGCTTCCTCG	352		

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NHL-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp667D1012) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 16 Row: P Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers
1. .3454
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="295.8 cR from top of Chr12 linkage group"
/clone="DKFZp667D1012"
/tissue_type="lymph node"
/clone_lib="5667 (synonym: hlno2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
1. .3454
/gene="DKFZp667D1012"
1. .464
/gene="DKFZp667D1012"
/note="unknown zinc finger protein"
/product="hypothetical protein"
/protein_id="CAD28468.1"
/db_xref="GI:19584344"
/translation="HTGERPYSCLKCEKTFGRHHLIRHQTHLHDKTSCSECGKNE
RCSHLASHORVHAEGKSGQVSGPTKRPRAPPVVKCHVTECKSFGRHHL
VRWLHTGKEKPPQCPKPCSKSGRKHLDRLHLLTHOGQSPRNSWDRTGSVF"
3288. .3293
/gene="DKFZp667D1012"
3306

FEATURES
Source
1. .596
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4052341"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NIH_MGC_58"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
BASE COUNT 286 a 82 c 107 g 121 t
ORIGIN

polyA_signal
polyA_site
BASE COUNT 1058 a 711 c 756 g 929 t
ORIGIN

Query Match 17.9%; Score 121; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 7.3e-55;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 556 ATCAAA 615
DB 450 ATCAAA 509
QY 616 AA 675
DB 510 AA 569
QY 676 A 676
DB 570 A 570

Query Match 18.0%; Score 122; DB 9; Length 3454;
Best Local Similarity 100.0%; Pred. No. 2.2e-55;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 555 CATCAAA 614
DB 3306 CATCAAA 3365
QY 615 AA 674
DB 3366 AA 3425
QY 675 AA 676
DB 3426 AA 3427

RESULT 3
BC034120
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BC034120 2776 bp mRNA linear ROD 08-JUL-2002
Mus musculus, RIKEN cDNA 2810429K11 gene, clone MGC:31645
IMAGE:4527027, mRNA, complete cds.
BC034120
BC034120.1 GI:21707638
MGC.
house mouse.

RESULT 2
BC008387
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BC008387 596 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, clone IMAGE:4052341, mRNA.
BC008387
BC008387.1 GI:14249997
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 596)
Strausberg, R.
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
NHL-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdexil1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 44 Row: j Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers

1..2776
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:31645 IMAGE:4527027"
/tissue_type="Eye, retina, mouse strain C57BL/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
205..2157

CDS

/codon_start=1
/product="RIKEN CDNA 2810429K11 gene"
/protein_id="AAH34120.1"
/db_xref="gi:21707639"
/db_xref="locusID:74199"
/translation="MGIVPTMKASVIEVLVLVLTGIHNSKTPKTKRPKLTVPQI
NCDVKGAGIINPEFVKPCAGQDPKYHVTGTYASYSVCGAAIHSGVLDNSGGKI
LVVRKAGSGYKGSYNGVOSLSLRWRESFVAESPKQGVAVPSLTYSKSTAA
KAGETTRAYEKPSIPGTTIOPTLTQAQATPAEVRHRSKPEASVYNSPRQVPG
HRSQMEEDVGKPGVLLDGSFVPEELSTQSEPVQDQPKCKDLSLDGSTSI
GKRFRIOKFLADYQALDGPAGPLVGVYQIGDNPATQFNKTHNSQDLKATIEK
ITORGGLSNVGRALSFVITKFFSKANGRGAPNAVVMVDGPTDRKEVSRAVES
GINVFPIVEGAARDIOHVPEGFASKYCRNTGFSFNQSWLSLHKTVQPLVLRV
CDTRLACSTCLNSADIGFVDSGSKNGTSENFVQFVANLSKEIISDTOTRGA
VOYTYEQLRSEFGDKYNSKADILSAIRRVGVSGGTGTGAIOVALBOLFKSKPNKR
KVMIIITDGRSYDDVRIPAMAAVQKGVITVAIGIAWAQAQDELEVMTHPAKDISFFVD
DFDNLKIAPRIQNICTEFNSQPN"

BASE COUNT 831 a 627 c 683 g 635 t

Query Match 17.9%; Score 121; DB 10; Length 2776;
Best Local Similarity 100.0%; Pred. No. 7.8e-55;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 556 ATCAAA 615
DB 2629 ATCAAA 2688
QY 616 AA 675
DB 2689 AA 2748
QY 676 A 676
DB 2749 A 2749

RESULT 4

LOCUS HSM803477 HSM803477 2830 bp mRNA linear PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp686E0316 (from clone DKFZp686E0316).
ACCESSION AL832170
VERSION AL832170.1 GI:21732714
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2830)
Wambutt, R., Heubner, D., Mewes, H.W., Well, B. and Wiemann, S.
Direct Submission
JOURNAL
COMMENT
Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
consortium of the German Genome Project.
This clone (DKFZp686E0316) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..2830
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686E0316"
/tissue_type="cDNA-collection"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiI + SfiIB"
/dev_stage="adult"
polyA_signal 2686..2691
polyA_site 2702
BASE COUNT 982 a 438 c 583 g 827 t

Query Match 17.9%; Score 121; DB 9; Length 2830;
Best Local Similarity 100.0%; Pred. No. 7.8e-55;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 556 ATCAAA 615
DB 2706 ATCAAA 2765
QY 616 AA 675
DB 2766 AA 2825
QY 676 A 676
DB 2826 A 2826

RESULT 5

LOCUS HSM803701 HSM803701 3383 bp mRNA linear PRI 10-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp667K1712 (from clone DKFZp667K1712).
ACCESSION AL832393
VERSION AL832393.1 GI:21732956
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3383)
Koehler, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.
Direct Submission
JOURNAL
COMMENT
Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp667K1712) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..3383
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="DKFZp667K1712"		BC034163.1		GI:21706753	
/tissue_type="lymph node"		MGC.		house mouse.	
DH10B: sites NotI + SalI		Mus musculus		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
/dev_stage="adult"		1 (bases 1 to 800)		Strausberg, R.	
3231..3236		Direct Submission		Submitted (02-JUL-2002)	
3250		Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		NIH-MGC Project URL: http://mgc.nci.nih.gov	
polyA_signal		1210 a		Contact: MGC help desk	
polyA_site		552 c		Email: cgapbs-re@mail.nih.gov	
BASE COUNT		552 c		663 g	
ORIGIN		17.9%; Score 121; DB 9; Length 3383;		Tissue Procurement: Gilbert Smith, Ph.D.	
Query Match		Best Local Similarity 100.0%; Pred. No. 7.9e-55;		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		556 ATCAAA		DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	
QY		556 ATCAAA		Web site: http://www-shgc.stanford.edu	
Db		3251 ATCAAA		Contact: (Dickson, Mark) mcd@paxil.stanford.edu	
QY		616 AA		Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
Db		3311 AA		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
QY		676 A 676		Series: IRAK Plate: 58 Row: 0 Column: 24	
Db		3371 A 3371		This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753829.	
RESULT 6		AX458534/c		Location/Qualifiers	
LOCUS		AX458534		1..800	
DEFINITION		Sequence 80 from Patent WO0246454.		/organism="Mus musculus"	
ACCESSION		AX458534		/db_xref="taxon:10090"	
VERSION		AX458534.1		/map="C2ECH II"	
KEYWORDS		synthetic construct.		/clone="MGC:36077 IMAGE:5065647"	
SOURCE		synthetic construct		/tissue_type="Mammary tumor metastasized to lung, Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."	
ORGANISM		artificial sequences.		/clone_lib="NCI_CGAP_Lu29"	
REFERENCE		1		/lab_host="DH10B"	
AUTHORS		Schacht, O.		/note="Vector: pCMV-SPORT6"	
TITLE		Diagnosis of diseases associated with angiogenesis		45..305	
JOURNAL		Patent: WO 0246454-A 80 13-JUN-2002;		/codon_start=1	
FEATURES		Epigenomics AG (DE)		/product="Fc receptor, IgE, high affinity I, gamma polypeptide"	
source		Location/Qualifiers		/protein_id="AAH34163.1"	
1..6486		1. .6486		/db_xref="GI:21706754"	
/organism="synthetic construct"		/db_xref="taxon:32630"		/db_xref="LocusID:14127"	
/note="chemically treated genomic DNA (Homo sapiens)"		1713 a		/translation="MISAVILFLLVLEQAAALGEPOLCYILDVFLYGVITLLYC RUKIQVRKAAIASREKADAVYTGILNRSQETETLKHEKPPQ"	
BASE COUNT		1713 a		183 c	
ORIGIN		176 c		124 g	
Query Match		17.9%; Score 121; DB 6; Length 6486;		334 a	
Best Local Similarity 100.0%; Pred. No. 8.1e-55;		Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		BASE COUNT	
QY		556 ATCAAA		ORIGIN	
Db		3483 ATCAAA		Query Match	
QY		616 AA		Best Local Similarity 100.0%; Pred. No. 2.6e-54;	
Db		3423 AA		Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY		676 A 676		557 TCACAAAAA	
Db		3363 A 3363		595 TCACAAAAA	
RESULT 7		BC034163		617 AAAAAAAAA	
LOCUS		BC034163		655 AAAAAAAAA	
DEFINITION		Mus musculus, Fc receptor, IgE, high affinity I, gamma polypeptide, clone MGC:36077 IMAGE:5065647, mRNA, complete cds.		RESULT 8	
ACCESSION		BC034163		AK026865	
				LOCUS	
				DEFINITION	
				AF182844 Homo sapiens VPS28 protein mRNA.	

ACCESSION AK026865
 VERSION AK026865.1 GI:10439824
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens adipose tissue cDNA to mRNA, clone_lib:ad
 clone:ADS001518.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Kawakami,T., Nouchi,S., Itoh,T., Shigeta,K., Senba,T.,
 Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
 Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
 Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
 Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1034)
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing; Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source Location/Qualifiers
 1..1034
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADS001518"
 /tissue_type="adipose tissue"
 /clone_lib="ad"
 /note="cloning vector pME18SFL3"
 misc_feature 1..1034
 /note="highly similar to AF182844 Homo sapiens VPS28
 protein mRNA"
 BASE COUNT 315 a 289 c 264 g 166 t
 ORIGIN
 Query Match 17.8%; Score 120; DB 9; Length 1034;
 Best Local Similarity 100.0%; Pred. No. 2.6e-54;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 557 TCAA 616
 Db 912 TCAA 971
 Qy 617 AA 676
 Db 972 AA 1031
 RESULT 9
 BC014547
 LOCUS BC014547 1191 bp mRNA linear PRI 26-SEP-2001
 DEFINITION Homo sapiens, similar to myoglobin, clone MGC:13548 IMAGE:4244838,
 mRNA, complete cds.
 ACCESSION BC014547
 VERSION BC014547.1 GI:15778932
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1191)
 AUTHORS Strausberg,R.

TITLE
JOURNALREMARK
COMMENT

Direct Submission
 Submitted (24-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amand@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 19 Row: g Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4885476.
 FEATURES
 source Location/Qualifiers
 1..1191
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:13548 IMAGE:4244838"
 /tissue_type="Skeletal Muscle"
 /clone_lib="NIH_MGC_81"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 77..541
 /codon_start=1
 /product="Similar to myoglobin"
 /protein_id="AAH14547.1"
 /db_xref="GI:15778933"
 /translation="MGLSDGKQLVNVGKVEADIPGHQGVLRIRLKGHPETLEKFD
 DKFKHLKSEDEKAKSADLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPV
 KYLEFISECIQVLIQKHGPDGADAGAMNKALELFKRKDMASNYKELGFG"
 BASE COUNT 357 a 278 c 314 g 242 t
 ORIGIN

CDS

Query Match 17.8%; Score 120; DB 9; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 2.6e-54;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 557 TCAA 616
 Db 1072 TCAA 1131
 Qy 617 AA 676
 Db 1132 AA 1191
 RESULT 10
 AF113689
 ID AF113689 standard; RNA; HUM; 1298 BP.
 AC AF113689;
 XX AF113689;
 SV AF113689.1
 XX
 DT 03-FEB-2000 (Rel. 62, Created)
 DT 09-MAY-2001 (Rel. 67, Last updated, Version 2)
 XX
 DE Homo sapiens clone FLB4701.
 XX
 KW
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 29 Row: m Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA qi: 4505790.

	BASE COUNT	480 a	250 c	279 g	310 t	ORIGIN
	Query Match		17.8%;	Score 120;	DB 9;	Length 1319;
	Best Local Similarity	100.0%;		Pred. No. 2.7e-34;		
	Matches 120;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	557	TCAA		AA		616
Db	1176	TCAAA		AA		1235
Qy	617	AA		AA		676
Db	1236	AA		AA		1295

RESULT-12	
BC016332	
LOCUS	
DEFINITION	linear mRNA
	1388 bp
	linear PRI 05-NOV-2001
	Homo sapiens, ubiquitin-conjugating enzyme E2 variant 2, clone
	MGC:23763 IMAGE:4109228, mRNA, complete cds.
ACCESSION	BC016332
VERSION	BC016332.1
KEYWORDS	GI:16740953
SOURCE	MGC.
ORGANISM	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1388)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (23-OCT-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgapbs-remail.nih.gov
	Tissue Procurement: ATCC
	cDNA Library Prepared by: CLONTECH Laboratories, Inc.
	cDNA Library Artaived by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadanesbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 32 Row: D Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..1388

/organism="Homo sapiens"

/db_xref="LocusID:7336"

/db_xref="taxon:9606"

/clone="MGC:23763 IMAGE:4109228"

/tissue_type="Bone marrow, acute myelogenous leukemia"

/clone_lib="NIH_MGC_55"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

47..484

/codon_start=1

/product="ubiquitin-conjugating enzyme E2 variant 2"

/protein_id="AAH16332.1"

/db_xref="GI:16740954"

/translation="MAVSTGVKPRNFRLLLEELLEGKGVGDGTVSGLEDDEDMTIT

RWCMIIIGPRTNENRIYLSKVECKGYPFAPPSVRFVTKINNINGINSSGMVDARS

IPVLAKWNSISYKIVQLQELRLMSKKNKLPPEPGQYNN"

592 a 179 c 271 g 346 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 17.8%; Score 120; DB 9; Length 1388;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAA 616

Db 1223 TCAA 1282

Qy 617 AA 676

Db 1283 AA 1342

RESULT 13

BC006018

LOCUS

DEFINITION

Mus musculus, MyoD family inhibitor, clone MGC:5939 IMAGE:3488088,

mRNA, complete cds.

ACCESSION

BC006018

VERSION

BC006018.1 GI:13543726

KEYWORDS

MGC.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1462)

Submitted (02-APR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

REMARK

COMMENT

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 7 Row: O Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754665.

FEATURES

Location/Qualifiers

1..1462

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="C57BL/6J"

/clone="MGC:5939 IMAGE:3488088"

/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months

old, gross tissue."

/clone_lib="NCI-CGAP_Mam5"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

26..766

/codon_start=1

/product="MyoD family inhibitor"

/protein_id="AAH06018.1"

/db_xref="GI:13543727"

/db_xref="LocusID:17240"

/translation="MSQVSGQCPSCRDAPHGVPSAALDPAQTMSLPLGLEVARSTHPV

ESSEHILSEFPEEAPSPHDSGLRAQALNSIDLDPTEAVTCQPGNPGCTPLLPNG

SSHDHLSSEFAGHAGNAGLGGSKAHRKLTQHPSLGSGAGRSKRSASQVPLQAG

ECCVHCILSCLECEFTLCNILLDCATCGSCSESDCLCCCGSGECADCDLPCLD

DCGIVDACCESADCLCIEHCCEGCLCFSS"

383 a 423 c 407 g 249 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 17.8%; Score 120; DB 10; Length 1462;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAA 616

Db 1335 TCAA 1394

Qy 617 AA 676

Db 1395 AA 1454

RESULT 14

BC032462

LOCUS

DEFINITION

Homo sapiens, hypothetical gene supported by BC017964, clone

MGC:40428 IMAGE:5197243, mRNA, complete cds.

ACCESSION

BC032462

VERSION

BC032462.1 GI:21595557

KEYWORDS

MGC.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1493)

Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lario, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 64 Row: h Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

FEATURES

source

1. 1493

/organism="Homo sapiens"

/db_xref="LocusID:144714"

/db_xref="taxon:9606"

/clone="MGC:40428 IMAGE:5197243"

/tissue_type="Brain, adult, 6 pooled whole brains"

/clone_lib="NIH_MGC_114"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

75..566

/codon_start=1

/product="hypothetical gene supported by BC017964"

/protein_id="AAH32462.1"

/db_xref="GI:21595558"

/translation="MYRAVAATKAWLGGVLRCWEPWCGRGIGKSRPRQPGGA GLGKAQKRSGREGRRPLGAAGLGLVSPQAGRHGRKKAPLAARPVTRHARPSQGR PLVLYLECSGDRGTVLRFQVTPTRFPKAPPAGWGVVVEAGLSDAEQEVASRISYFEN FAI"

BASE COUNT

403 a 371 c 401 g 318 t

ORIGIN

Query Match 17.8%; Score:120; DB 9; Length 1493;
Best Local Similarity 100.0%; Pred. No. 2.7e-54;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAAA 616
|||||
Db 1356 TCAAA 1415
|||||
Qy 617 AAA 676
|||||
Db 1416 AAA 1475
|||||

RESULT 15

XLNAKATP

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

XLNAKATP

X.laeviss mRNA for Na,K-ATPase gamma subunit.

Y11587

Y11587.1

GI:2052283

ATPase; gamma subunit; gamma subunit of sodium potassium ATPase; NA+, K+-ATPase.
Xenopus laevis.

Xenopus laevis

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.
1 (bases 1 to 1500)
Beguin, P., Wang, X., Firsov, D., Puoti, A., Claeys, D., Horisberger, J.D. and Geering, K.

The gamma subunit is a specific component of the Na,K-ATPase and modulates its transport function

JOURNAL

EMBO J. 16 (14), 4250-4260 (1997)

MEDLINE

97392454

PUBMED

9250668

REFERENCE

2 (bases 1 to 1500)

AUTHORS

Beguin, P.

TITLE

Direct Submission

JOURNAL

Submitted (26-FEB-1997) P. Beguin, University of Lausanne, Institute of Pharmacology and Toxicology, Rue du Bugnon 27, 1005 Lausanne, Switzerland

FEATURES

source

1. 1500

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/cell_line="A6"

/tissue_type="kidney"

63..248

/codon_start=1

/product="gamma subunit of Na,K-ATPase"

/protein_id="CAA72326.1"

/db_xref="GI:2052284"

/db_xref="SWISS-PROT:O13001"

/translation="MADAQDDMSQMDKFTYDYETIRKGLIFALIAFVVGMLIFSG RFRGGRKKLRALNDDM"

BASE COUNT

559 a 249 c 260 g 432 t

ORIGIN

Query Match 17.8%; Score:120; DB 5; Length 1500;
Best Local Similarity 100.0%; Pred. No. 2.7e-54;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAAA 616
|||||
Db 1349 TCAAA 1403
|||||
Qy 617 AAA 676
|||||
Db 1409 AAA 1463
|||||

Search completed: June 16, 2003, 07:55:55

Job time : 2000.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 03:58:29 ; Search time 222.269 seconds
(without alignments)
5849.143 Million cell updates/sec

Title: us-10-090-035-1
Perfect score: 676
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Genesec101002.*
1: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/genesec/genesecq-n-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121	17.9	381	23	ABV58662 Human prostate exp
2	121	17.9	1119	21	AAC60033 Human secreted pro
3	121	17.9	2645	22	AAS31262 Human cDNA encodin
4	121	17.9	2645	24	AB066586 Human polynucleoti
5	121	17.9	6486	24	AB067050 Human anglogenesis
6	120	17.8	439	23	ABV50206 Human prostate exp
7	120	17.8	452	23	ABV59282 Human prostate exp
8	120	17.8	484	23	ABV58008 Human prostate exp
9	120	17.8	516	23	ABV55889 Human prostate exp

10	120	17.8	544	23	ABV56624 Human prostate exp
11	120	17.8	579	23	ABV58690 DNA encoding human
12	120	17.8	804	22	AAS33148 Human secreted pro
13	120	17.8	1097	22	AAD08201 Human secreted pro
14	120	17.8	1097	22	AAD05430 DNA encoding novel
15	120	17.8	1194	23	ABK43496 Human polynucleoti
16	120	17.8	1204	22	AAI63823 Human secreted pro
17	120	17.8	1457	21	AAA61287 Human secreted pro
18	120	17.8	1985	21	AAC59406 Mouse neuronal PAS
19	120	17.8	2082	19	AAV41257 Human cDNA clone B
20	120	17.8	2178	22	AAF98380 Human immune syste
21	120	17.8	6171	24	ABL32788 Chemically treated
22	120	17.8	6577	24	ABL70561 Human immune syste
23	120	17.8	6577	24	ABL33356 Human immune syste
24	120	17.8	6577	24	AAS61221 Chemically treated
25	120	17.8	6794	24	ABL70219 Human gene regulat
26	120	17.8	6794	24	AAS61174 Human gene regulat
27	120	17.8	6794	24	ABK31264 Signal transductio
28	120	17.8	8899	24	ABK40074 Human chemically p
29	120	17.8	8899	24	ABL32837 Human immune syste
30	120	17.8	8946	24	ABL32911 Chemically pretrea
31	120	17.8	15832	22	AAS45399 Human immune syste
32	120	17.8	15832	24	ABL33343 DNA transcription
33	120	17.8	15832	24	ABK28244 Human anglogenesis
34	120	17.8	40324	24	ABQ67149 Human nervous syst
35	119	17.6	165	22	ABA17044 Human ovarian and
36	119	17.6	165	22	AAS39812 Genomic sequence #
37	119	17.6	165	22	AAD16672 Human pancreatic f
38	119	17.6	165	22	AAS29235 Human reproductive
39	119	17.6	165	22	AAI04803 Human reproductive
40	119	17.6	165	22	AAI07127 Human immune/haema
41	119	17.6	165	22	AAK73944 Human immune/haema
42	119	17.6	165	22	AAK78761 Human immune/haema
43	119	17.6	165	22	AAK79574 Human immune/haema
44	119	17.6	165	22	AAK81312 Human immune/haema
45	119	17.6	165	22	AAK81312 Human immune/haema

ALIGNMENTS

RESULT 1
ABV58662
ID ABV58662 standard; cDNA; 381 BP.
XX AC ABV58662;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 58653.
DE KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX WO200160860-A2.
PN PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;

—

PN	WO200155368-A1.		
XX	02-AUG-2001.		
PD	17-JAN-2001; 2001WO-US01348.		
XX	31-JAN-2000; 2000US-0179065.		
PF	04-FEB-2000; 2000US-0180628.		
XX	24-FEB-2000; 2000US-0184664.		
	02-MAR-2000; 2000US-0186350.		
	16-MAR-2000; 2000US-0198874.		
	17-MAR-2000; 2000US-0190076.		
	18-APR-2000; 2000US-0198123.		
	19-MAY-2000; 2000US-0205515.		
	07-JUN-2000; 2000US-0209467.		
	28-JUN-2000; 2000US-0214886.		
	30-JUN-2000; 2000US-0215135.		
	07-JUL-2000; 2000US-0216647.		
	07-JUL-2000; 2000US-0216880.		
	11-JUL-2000; 2000US-0217487.		
	11-JUL-2000; 2000US-0217496.		
	14-JUL-2000; 2000US-0218220.		
	26-JUL-2000; 2000US-0220963.		
	26-JUL-2000; 2000US-0220964.		
	14-AUG-2000; 2000US-0224518.		
	14-AUG-2000; 2000US-0225213.		
	14-AUG-2000; 2000US-0225214.		
	14-AUG-2000; 2000US-0225266.		
	14-AUG-2000; 2000US-0225267.		
	14-AUG-2000; 2000US-0225268.		
	14-AUG-2000; 2000US-0225270.		
	14-AUG-2000; 2000US-0225447.		
	14-AUG-2000; 2000US-0225757.		
	14-AUG-2000; 2000US-0225758.		
	14-AUG-2000; 2000US-0225759.		
	18-AUG-2000; 2000US-0226279.		
	22-AUG-2000; 2000US-0226681.		
	22-AUG-2000; 2000US-0226868.		
	22-AUG-2000; 2000US-0227182.		
	23-AUG-2000; 2000US-0227009.		
	30-AUG-2000; 2000US-0228924.		
	01-SEP-2000; 2000US-0229287.		
	01-SEP-2000; 2000US-0229343.		
	01-SEP-2000; 2000US-0229344.		
	01-SEP-2000; 2000US-0229345.		
	05-SEP-2000; 2000US-0229509.		
	05-SEP-2000; 2000US-0229513.		
	06-SEP-2000; 2000US-0230437.		
	06-SEP-2000; 2000US-0230438.		
	08-SEP-2000; 2000US-0231142.		
	08-SEP-2000; 2000US-0231243.		
	08-SEP-2000; 2000US-0231244.		
	08-SEP-2000; 2000US-0231413.		
	08-SEP-2000; 2000US-0231414.		
	08-SEP-2000; 2000US-0232080.		
	08-SEP-2000; 2000US-0232081.		
	12-SEP-2000; 2000US-0231968.		
	14-SEP-2000; 2000US-0232397.		
	14-SEP-2000; 2000US-0232398.		
	14-SEP-2000; 2000US-0232399.		
	14-SEP-2000; 2000US-0232400.		
	14-SEP-2000; 2000US-0232401.		
	14-SEP-2000; 2000US-0233063.		
	14-SEP-2000; 2000US-0233064.		
	21-SEP-2000; 2000US-0233065.		
	21-SEP-2000; 2000US-0234223.		
	21-SEP-2000; 2000US-0234274.		
	25-SEP-2000; 2000US-0234997.		
	25-SEP-2000; 2000US-0234998.		
	26-SEP-2000; 2000US-0235484.		
	27-SEP-2000; 2000US-0235834.		
	27-SEP-2000; 2000US-0235836.		
PR	29-SEP-2000; 2000US-0236327.		
PR	29-SEP-2000; 2000US-0236367.		
PR	29-SEP-2000; 2000US-0236368.		
PR	29-SEP-2000; 2000US-0236369.		
PR	29-SEP-2000; 2000US-0236370.		
PR	02-OCT-2000; 2000US-0236802.		
PR	02-OCT-2000; 2000US-0237037.		
PR	02-OCT-2000; 2000US-0237038.		
PR	02-OCT-2000; 2000US-0237039.		
PR	02-OCT-2000; 2000US-0237040.		
PR	13-OCT-2000; 2000US-0239335.		
PR	13-OCT-2000; 2000US-0239337.		
PR	20-OCT-2000; 2000US-0240960.		
PR	20-OCT-2000; 2000US-0241221.		
PR	20-OCT-2000; 2000US-0241785.		
PR	20-OCT-2000; 2000US-0241786.		
PR	20-OCT-2000; 2000US-0241787.		
PR	20-OCT-2000; 2000US-0241808.		
PR	20-OCT-2000; 2000US-0241809.		
PR	01-NOV-2000; 2000US-0241826.		
PR	01-NOV-2000; 2000US-0244617.		
PR	08-NOV-2000; 2000US-0246474.		
PR	08-NOV-2000; 2000US-0246475.		
PR	08-NOV-2000; 2000US-0246476.		
PR	08-NOV-2000; 2000US-0246477.		
PR	08-NOV-2000; 2000US-0246478.		
PR	08-NOV-2000; 2000US-0246523.		
PR	08-NOV-2000; 2000US-0246524.		
PR	08-NOV-2000; 2000US-0246525.		
PR	08-NOV-2000; 2000US-0246526.		
PR	08-NOV-2000; 2000US-0246527.		
PR	08-NOV-2000; 2000US-0246528.		
PR	08-NOV-2000; 2000US-0246532.		
PR	08-NOV-2000; 2000US-0246610.		
PR	08-NOV-2000; 2000US-0246611.		
PR	08-NOV-2000; 2000US-0246613.		
PR	17-NOV-2000; 2000US-0249207.		
PR	17-NOV-2000; 2000US-0249208.		
PR	17-NOV-2000; 2000US-0249209.		
PR	17-NOV-2000; 2000US-0249210.		
PR	17-NOV-2000; 2000US-0249211.		
PR	17-NOV-2000; 2000US-0249212.		
PR	17-NOV-2000; 2000US-0249213.		
PR	17-NOV-2000; 2000US-0249214.		
PR	17-NOV-2000; 2000US-0249215.		
PR	17-NOV-2000; 2000US-0249216.		
PR	17-NOV-2000; 2000US-0249217.		
PR	17-NOV-2000; 2000US-0249218.		
PR	17-NOV-2000; 2000US-0249244.		
PR	17-NOV-2000; 2000US-0249245.		
PR	17-NOV-2000; 2000US-0249264.		
PR	17-NOV-2000; 2000US-0249265.		
PR	17-NOV-2000; 2000US-0249297.		
PR	17-NOV-2000; 2000US-0249299.		
PR	17-NOV-2000; 2000US-0249300.		
PR	01-DEC-2000; 2000US-0250160.		
PR	01-DEC-2000; 2000US-0250391.		
PR	05-DEC-2000; 2000US-0251030.		
PR	05-DEC-2000; 2000US-0251988.		
PR	05-DEC-2000; 2000US-0256719.		
PR	06-DEC-2000; 2000US-0251479.		
PR	08-DEC-2000; 2000US-0251856.		
PR	08-DEC-2000; 2000US-0251868.		
PR	08-DEC-2000; 2000US-0251869.		
PR	08-DEC-2000; 2000US-0251989.		
PR	08-DEC-2000; 2000US-0251990.		
PR	11-DEC-2000; 2000US-0254097.		
PR	05-JAN-2001; 2001US-0259678.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA, Barash SC, Ruben SM;		
XX			
PI			

[illegible]

CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870.

XX Sequence 2645 BP; 827 A; 506 C; 513 G; 799 T; 0 other;

Query Match 17.9%; Score 121; DB 24; Length 2645;
 Best Local Similarity 100.0%; Pred. No. 1.7e-28;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ATCAAA 615
 DB 2522 ATCAAA 2581

QY 616 AA 675
 DB 2582 AA 2641

QY 676 A 676
 DB 2642 A 2642

RESULT 5

ABQ67050/c
 ID ABQ67050 standard; DNA; 6486 BP.

XX AC ABQ67050;

XX DT 28-AUG-2002 (first entry)

XX DE Human angiogenesis associated polynucleotide SEQ ID NO 80.

XX KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
 KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
 KW macular degeneration; inflammatory bowel disease; Crohn's disease;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiarteriosclerotic; ds.

XX OS Homo sapiens.

XX PN WO200246454-A2.

XX PD 13-JUN-2002.

XX PF 06-DEC-2001; 2001WO-EP14320.

XX PR 06-DEC-2000; 2000DE-1061338.

XX PA (EPIG-) EPIGENOMICS AG.

XX PT Schacht O;

XX DR WPI; 2002-500450/53.

XX PT New nucleic acid fragments from chemically treated
 PT angiogenesis-associated genes, useful for determining methylation
 status, e.g. in diagnosis or treatment of cancer

XX PS Claim 1; SEQ ID NO 80; 41pp + Sequence Listing; German.

XX

CC The invention relates to a nucleic acid (I) comprising a segment of 18
 CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
 CC related oligomers, are used to evaluate the methylation status and/or
 CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
 CC diagnosis and treatment of eye diseases, proliferative retinopathy,
 CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 CC diabetic retinopathy, macular degeneration caused by neovascularisation,
 CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 CC Crohn's disease.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 6486 BP; 1713 A; 176 C; 1443 G; 3154 T; 0 other;

Query Match 17.9%; Score 121; DB 24; Length 6486;
 Best Local Similarity 100.0%; Pred. No. 1.3e-28;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ATCAAA 615
 DB 3483 ATCAAA 3424

QY 616 AA 675
 DB 3423 AA 3364

QY 676 A 676

DB 3363 A 3363

RESULT 6

ABV50206

ID ABV50206 standard; cDNA; 439 BP.

XX AC ABV50206;

XX DT 17-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 50197.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer

XX PS Claim 1; Page 9776; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 439 BP; 179 A; 117 C; 69 G; 69 T; 5 other;
Query Match 17.8%; Score 120; DB 23; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.1e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 113 TCAA 172
QY 617 AA 676
DB 173 AA 232
RESULT 7
ABV59282
ID ABV59282 standard; cDNA; 452 BP.
XX AC ABV59282;
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 59273.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX Claim 1; Page 11362; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 452 BP; 199 A; 91 C; 72 G; 89 T; 1 other;
Query Match 17.8%; Score 120; DB 23; Length 452;
Best Local Similarity 100.0%; Pred. No. 6e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 243 TCAA 302
QY 617 AA 676
DB 303 AA 362
RESULT 8
ABV58008
ID ABV58008 standard; cDNA; 484 BP.
XX AC ABV58008;
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 57999.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX Claim 1; Page 11145; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 544 BP; 220 A; 141 C; 97 G; 84 T; 2 other;
SQ
Query Match 17.8%; Score 120; DB 23; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.7e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 123 TCAA 182
QY 617 AA 676
DB 183 AA 242
RESULT 11
ID ABV58690 standard; cDNA; 579 BP.
XX
XX AC ABV58690;
XX
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 58681.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX
XX Claim 1; Page 11260-11261; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 579 BP; 269 A; 106 C; 82 G; 121 T; 1 other;
SQ
Query Match 17.8%; Score 120; DB 23; Length 579;
Best Local Similarity 100.0%; Pred. No. 5.6e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 168 TCAA 227
QY 617 AA 676
DB 228 AA 287
RESULT 12
AAS33148
ID AAS33148 standard; cDNA; 804 BP.
XX
XX AC AAS33148;
XX
XX 04-DEC-2001 (first entry)
XX
XX DNA encoding human secreted protein, Seq ID No 107.
XX
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX multiple sclerosis; cancer; hyperproliferative disorder; infection;
XX Gaucher's disease; neurological disease; cerebrovascular disorder;
XX thrombosis; wound healing; ss.
XX
XX Homo sapiens.
XX
XX WO200155326-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01347.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451931/48.
XX P-PSDB; AAU20439.
XX
XX New nucleic acids and polypeptides, useful for diagnosing, preventing
XX or treating medical conditions -
XX
XX Claim 1; SEQ ID No 107; 753pp; English.
XX
XX The invention relates to novel isolated nucleic acid molecules (I)
XX encoding human secreted proteins (II). (I) and (II) are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits, in
XX goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used
XX the prevention, treatment and diagnosis of diseases associated with
XX inappropriate expression of secreted proteins. (I) and complementary
XX sequences may also be used as DNA probes in diagnostic assays (e.g.
XX polynucleotide chain reactions (PCR)) to detect and quantitate the presence
XX of similar nucleic acid sequences in samples, and so which patients may
XX be in need of restorative therapy. (II) may also be used as antigens in
XX the production of antibodies and in assays to identify modulators

XX DT 17-JUL-2001 (first entry)

DE DE Human secreted protein-encoding gene 7 cDNA clone HCNV12, SEQ ID NO:52.

XX XX Human; secreted protein; proliferative disorder; cancer; tumour;

XX KW foetal abnormality; developmental abnormality; haematopoietic disorder;

XX KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

XX KW inflammation; allergy; neurological disorder; Alzheimer's disease;

XX KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

XX KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

XX KW cardiovascular disorder; angioinetic disorder; kidney disorder;

XX KW gastrointestinal disorder; pregnancy-related disorder;

XX KW endocrine disorder; infection; wound healing; vulnerability;

XX KW cell culture; chemotaxis; food additive; gene therapy;

XX KW binding partner identification; ss.

OS Homo sapiens.

XX XX Key Location/Qualifiers

XX CDS 193..378

FT /tag- a

FT /product- "Human secreted protein"

FT sig_peptide 193..264

FT /tag- b

FT mat_peptide 265..375

FT /tag- c

FT /product- "Mature human secreted protein"

XX XX W0200134623-AL.

XX XX 17-MAY-2001.

XX XX 01-NOV-2000; 2000WO-US30037.

XX XX 05-NOV-1999; 99US-0163577.

XX XX 30-JUN-2000; 2000US-0215137.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Ruben SM, Komatsoulis GA, Moore PA;

XX WPI; 2001-316490/33.

XX P-PSDB; AAE01587.

XX Nucleic acids encoding 29 human secreted polypeptides, useful for

XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's

XX disease and diabetic retinopathy -

XX Claim 1; Page 446; 535pp; English.

XX AAD05389-AAD05473 represent cDNAs corresponding to 29 human secreted

XX protein genes, and AAE01546-AAE01630 represent the proteins they encode.

XX AAE01631-AAE01660 represent human secreted protein fragments or variants.

XX The secreted proteins and their genes are useful for preventing, gene

XX treating or ameliorating medical conditions, e.g., by protein or gene

XX therapy. Pathological conditions can be diagnosed by determining the

XX amount of the new protein in a sample or by determining the presence of

XX mutations in the new genes. Specific uses are described for each of the

XX 29 genes, based on the tissues in which they are most highly expressed,

XX and include developing products for the diagnosis or treatment of

XX proliferative disorders, cancer, tumours, foetal and developmental

XX abnormalities, haematopoietic disorders, diseases of the immune system,

XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

XX allergies, neurological disorders (e.g., Alzheimer's disease,

XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,

XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

XX cardiovascular disorders, angioinetic disorders, kidney disorders,

XX gastrointestinal disorders, pregnancy-related disorders, endocrine

XX disorders, and infections. The proteins can also be used to aid wound

XX healing and epithelial cell proliferation, to prevent skin aging due to

XX sunburn, to maintain organs before transplantation, for supporting cell

XX culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked

CC immunosorbent assay (ELISA). The present sequence represents a human

XX secreted protein-encoding cDNA of the invention.

SQ Sequence 1097 BP; 458 A; 182 C; 231 G; 226 T; 0 other;

Query Match 17.8%; Score 120; DB 22; Length 1097;

Best Local Similarity 100.0%; Pred. No. 4.5e-28;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 TCAA 616

DB 977 TCAA 1036

QY 617 AA 676

DB 1037 AA 1096

RESULT 15

ABK43496

ID ABK43496 standard; cDNA; 1194 BP.

XX AC ABK43496;

XX DT 05-JUN-2002 (first entry)

DE DNA encoding novel central nervous system protein #76.

XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;

XX KW hyperproliferative disorder; neoplasm; cardiovascular disorder;

XX KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;

XX KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;

XX KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;

XX KW adenocarcinoma; reproductive system disorder; testicular feminisation;

XX KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;

XX KW respiratory disorder; renal disorder; kidney failure; blood disorder;

XX KW myocardial infarction; wound healing; cell proliferation; skin aging;

XX KW food additive; food preservative; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN W0200155318-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01332.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216847.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226888.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-581633/65.

P-PSDB; AAU87166.

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -

Claim 1; SEQ ID No 86; 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (II) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial

CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 17.8%; Score 120; DB 23; Length 1194;
 Best Local Similarity 100.0%; Pred. No. 4.4e-28;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 TCACAAAAA 616
 Db 1070 TCACAAAAA 1129
 QY 617 AAAAAAA 676
 Db 1130 AAAAAAA 1189

Search completed: June 16, 2003, 06:54:02
 Job time : 230.769 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 06:35:31 ; Search time 47.0496 Seconds
(without alignments)
4406.275 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 acccacgctccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	17.8	2082	2	US-08-785-310A-2
2	119	17.6	3275	4	US-09-370-838-151
3	118	17.5	144	1	US-08-702-344-26
4	118	17.5	240	1	US-08-628-417-6
5	118	17.5	1051	4	US-09-245-041-10
6	118	17.5	1447	4	US-09-443-041A-27
7	118	17.5	1798	4	US-09-797-906-1
8	118	17.5	2447	2	US-09-014-969-14
9	118	17.5	2674	4	US-09-817-180-1
10	117	17.3	117	1	US-08-702-344-3
11	117	17.3	1733	3	US-09-073-569-1
12	116	17.2	1813	4	US-09-071-224-3
13	114	16.9	2269	4	US-09-394-645-1
14	114	16.9	2269	4	US-09-243-560B-1
15	113	16.7	2246	4	US-09-363-708-3
16	111	16.4	2634	4	US-09-463-238-3
17	109	16.1	111	4	US-09-297-535-23
18	109	16.1	121	4	US-09-297-535-20
19	109	16.1	2186	4	US-09-360-545-66
c 20	108	16.0	9589	1	US-07-925-695-1
c 21	108	16.0	9589	1	US-07-925-695-2
22	106	15.7	2184	4	US-08-955-918C-1
23	106	15.7	2184	4	US-08-697-786A-1
24	105	15.5	1872	4	US-09-801-052-1
25	104	15.4	1066	1	US-08-157-101A-4
26	103	15.2	1882	4	US-09-370-253-1
27	101	14.9	140	1	US-08-628-417-5

28	100	14.8	578	4	US-09-602-877A-95	Sequence 95, Appl
29	99	14.6	1474	4	US-08-821-994-64	Sequence 64, Appl
30	99	14.6	6671	1	US-08-280-443-1	Sequence 1, Appl
31	99	14.6	6671	1	US-08-457-459-1	Sequence 1, Appl
32	99	14.6	6671	1	US-08-535-678-1	Sequence 1, Appl
33	99	14.6	6671	5	PCT-US95-02275-1	Sequence 1, Appl
34	98	14.5	98	1	US-08-088-658-42	Sequence 42, Appl
35	98	14.5	98	2	US-08-471-907A-42	Sequence 42, Appl
36	98	14.5	1454	4	US-09-372-422A-19	Sequence 19, Appl
37	96	14.2	1117	4	US-09-247-373B-33	Sequence 33, Appl
38	95	14.1	2806	4	US-09-653-839-9	Sequence 9, Appl
39	93	13.8	1098	3	US-09-248-335-35	Sequence 35, Appl
40	92	13.6	1582	3	US-08-545-196B-10	Sequence 10, Appl
41	92	13.6	1582	3	US-08-545-196B-12	Sequence 12, Appl
42	91	13.5	2323	4	US-09-149-476-24	Sequence 24, Appl
43	91	13.5	2671	6	5168051-9	Patent No. 5168051
44	90	13.3	2665	4	US-08-971-089-5	Sequence 5, Appl
45	89	13.2	1193	4	US-09-372-422A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-785-310A-2
; Sequence 2, Application US/08785310A
; Patent No. 5840532
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; TITLE OF INVENTION: Neuronal PAS Domain Protein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,310A
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-785-310A-2

Query Match 17.8%; Score 120; DB 2; Length 2082;
Best Local Similarity 100.0%; Pred. No. 1.8e-35;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 1958 TCNAAA 2017
QY 617 AA 676

[illegible]

RESULT 5

```

US-09-245-041-10
, Sequence 10, Application US/09245041
, Patent No. 6274339
, GENERAL INFORMATION:
, APPLICANT: Moore, K.
, APPLICANT: Nogle, D.
, TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
, TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
, FILE REFERENCE: 7853-136
, CURRENT APPLICATION NUMBER: US/09/245,041
, CURRENT FILING DATE: 1999-02-05
, EARLIER APPLICATION NUMBER: 60/093,630
, EARLIER FILING DATE: 1998-07-21
, EARLIER APPLICATION NUMBER: 60/104,978
, EARLIER FILING DATE: 1998-10-20
, NUMBER OF SEQ ID NOS: 131
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 10
, LENGTH: 1051
, TYPE: DNA
, ORGANISM: Mus musculus
US-09-245-041-10

```

	Query Match	Best Local Similarity	Score 118; DB 4;	Length 1051;
	Matches 118;	Conservative	Mismatches 0;	Indels 0;
	Gaps 0;			
OY	559	AAA 	17.5%;	618
Ddb	10	AAA 	100.0%; Pred. No. 1.e-34;	69
OY	619	AAA 	17.5%;	676
Ddb	70	AAA 	100.0%; Pred. No. 1.e-34;	127

RESULT 6

```

US-09-443-041A-27
; Sequence 27, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shep, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-443-041A-27

Query Match 17.5%; Score 118; DB 4; Length 1447;
Best Local Similarity 100.0%; Pred. No. 1e-34;

```

RESULT 7

```

US-09-797-906-1
; Sequence 1, Application US/09797906
; Patent No. 6323188
;
; GENERAL INFORMATION:
; APPLICANT: ZIANGHE YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US$ THEREOF
; FILE REFERENCE: CLO01151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1798
; TYPE: DNA
; ORGANISM: Human
US-09-797-906-1

```

[illegible]

RESULTS

US-09-014-969-14
; Sequence 14, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-014-969-14

Query Match 17.5%; Score 118; DB 2; Length 2447;
Best Local Similarity 100.0%; Pred. No. 9.3e-35;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 2260 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2319

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
Db 2320 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2377

RESULT 9
US-09-817-180-1
; Sequence 1, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CL001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
; US-09-817-180-1

Query Match 17.5%; Score 118; DB 4; Length 2674;
Best Local Similarity 100.0%; Pred. No. 9.1e-35;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 2547 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2606

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
Db 2607 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2664

RESULT 10
US-08-702-344-3
; Sequence 3, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa

; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-702-344-3

Query Match 17.3%; Score 117; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.1e-34;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 675
Db 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 117

RESULT 11
US-09-073-569-1
; Sequence 1, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 34...1344
; OTHER INFORMATION:
;
US-09-073-569-1

Query Match 17.3% Score 117; DB 3; Length 1733;
Best Local Similarity 100.0%; Pred. No. 2.3e-34;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 558 CAAAAA... 617
Db 1607 CAAAAA... 1666
Qy 618 AAAAAA... 674
Db 1667 AAAAAA... 1723

RESULT 12
US-09-071-224-3
; Sequence 3, Application US/09071224
; Patent No. 6271343
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Presnell, Scott R.
; APPLICANT: Jelmsberg, Anna C.
; APPLICANT: Gilbert, Teresa
; APPLICANT: Foster, Donald C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Lehner, Joyce M.
; TITLE OF INVENTION: MAMMALIAN ZCYTORS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics
; STREET: 1201 Eastlake Ave East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,224
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1813 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 88...1362
; OTHER INFORMATION:
;
US-09-071-224-3

Query Match 17.2% Score 116; DB 4; Length 1813;
Best Local Similarity 100.0%; Pred. No. 5.2e-34;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 558 CAAAAA... 617
Db 1687 CAAAAA... 1746
Qy 618 AAAAAA... 673
Db 1747 AAAAAA... 1802

RESULT 13
US-09-394-645-1
; Sequence 1, Application US/09394645
; Patent No. 6380371
; GENERAL INFORMATION:
; APPLICANT: Sassetti, Christopher M.
; APPLICANT: Rosen, Steven R.
; TITLE OF INVENTION: Endoglycan: A NOVEL PROTEIN HAVING SELECTIN
; TITLE OF INVENTION: LIGAND AND CHEMOKINE PRESENTATION ACTIVITY
; FILE REFERENCE: 6510-122US1
; CURRENT APPLICATION NUMBER: US/09/394,645
; CURRENT FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: human
;
US-09-394-645-1

Query Match 16.9% Score 114; DB 4; Length 2269;
Best Local Similarity 100.0%; Pred. No. 2.6e-33;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 559 AAAAAA... 618
Db 2156 AAAAAA... 2215
Qy 619 AAAAAA... 672
Db 2216 AAAAAA... 2269

RESULT 14
US-09-243-560B-1
; Sequence 1, Application US/09243560B
; Patent No. 6395882
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sassetti, Christopher M.
; TITLE OF INVENTION: No. 6395882el Selectin Ligands
```

```
FILE REFERENCE: UCAL097US1
CURRENT APPLICATION NUMBER: US/09/243,560B
CURRENT FILING DATE: 1999-02-03
PRIOR APPLICATION NUMBER: US 60/074,389
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2269
TYPE: DNA
ORGANISM: Homo sapiens
US-09-243-560B-1

Query Match      16.9%; Score 114; DB 4; Length 2269;
Best Local Similarity 100.0%; Pred. No. 2.6e-33;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2156 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2215

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 672
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2216 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2269

RESULT 15
US-09-363-708-3
; Sequence 3, Application US/09363708
; Patent No. 6399747
; GENERAL INFORMATION:
; APPLICANT: Schmandt, et al.
; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 01017/34451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; DESCRIPTION: /desc = "mouse PAL cDNA"
US-09-363-708-3

Query Match      16.7%; Score 113; DB 4; Length 2246;
Best Local Similarity 100.0%; Pred. No. 6.1e-33;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 617
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 2134 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2193
QY 618 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 670
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2194 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2246

Search completed: June 16, 2003, 08:47:01
Job time : 54.0496 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 07:57:31 ; Search time 149.802 Seconds
(without alignments)
6534.570 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 acccagcgctcgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1029858 seqs, 724030393 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	676	100.0	676	9	US-10-090-035-1
2	313	46.3	574	9	US-10-090-035-3
3	313	46.3	570	9	US-10-090-035-5
4	261	38.6	580	9	US-10-090-035-7
5	188	27.8	529	9	US-10-090-035-9
6	169	25.0	524	9	US-10-090-035-17
7	121	17.9	2645	9	US-10-125-540-76
8	121	17.9	2645	10	US-09-764-870-76
9	120	17.8	1097	9	US-09-832-129-21
10	120	17.8	1457	9	US-09-948-820-38
c 11	120	17.8	15832	9	US-10-239-676-106
c 12	119	17.6	160	10	US-09-960-352-11058
13	119	17.6	165	9	US-09-764-872-709
c 14	119	17.6	165	9	US-10-091-483-344
15	119	17.6	165	9	US-09-764-891-7491
16	119	17.6	165	9	US-09-764-891-9815
c 17	119	17.6	165	10	US-09-764-846-344
c 18	119	17.6	181	10	US-09-960-352-1301
c 19	119	17.6	217	10	US-09-960-352-6539

20	119	17.6	409	10	US-09-822-849A-37	Sequence 37, Appl
21	119	17.6	495	10	US-09-878-574-4765	Sequence 4765, Ap
22	119	17.6	569	9	US-10-012-542-44	Sequence 44, Appl
23	119	17.6	674	9	US-10-076-622-465	Sequence 465, App
24	119	17.6	674	9	US-09-551-621-465	Sequence 21, Appl
25	119	17.6	674	10	US-09-825-301-21	Sequence 465, App
26	119	17.6	674	10	US-09-604-287A-465	Sequence 465, App
27	119	17.6	674	12	US-10-007-805-465	Sequence 26, Appl
c 28	119	17.6	681	10	US-09-822-830A-26	Sequence 161, Appl
29	119	17.6	758	10	US-09-764-853-161	Sequence 67, Appl
30	119	17.6	831	10	US-09-729-674-67	Sequence 212, Appl
31	119	17.6	1985	9	US-09-905-291A-212	Sequence 212, App
32	119	17.6	1985	9	US-09-902-853-212	Sequence 212, App
33	119	17.6	1985	9	US-09-907-824-212	Sequence 212, App
34	119	17.6	1985	9	US-09-907-841-212	Sequence 212, App
35	119	17.6	1985	9	US-09-904-011-212	Sequence 212, App
36	119	17.6	1985	9	US-10-174-590-27	Sequence 27, Appl
37	119	17.6	1985	9	US-10-176-758-27	Sequence 27, Appl
38	119	17.6	1985	9	US-10-175-737-27	Sequence 27, Appl
39	119	17.6	1985	9	US-09-906-742-212	Sequence 212, App
40	119	17.6	1985	9	US-10-173-706-27	Sequence 27, Appl
41	119	17.6	1985	9	US-10-175-738-27	Sequence 27, Appl
42	119	17.6	1985	9	US-10-175-752-27	Sequence 27, Appl
43	119	17.6	1985	9	US-10-176-482-27	Sequence 27, Appl
44	119	17.6	1985	9	US-10-176-757-27	Sequence 27, Appl
45	119	17.6	1985	9	US-10-176-913-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-10-090-035-1
; Sequence 1, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(367)
US-10-090-035-1

Query Match	100.0%	Score 676;	DB 9;	Length 676;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 676;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACCCACGCGTCCGCCACCGCTCCGAGCAATCCACAAAGCACCTTGAAGGACCACTGC	60	
Db	1	ACCCACGCGTCCGCCACCGCTCCGAGCAATCCACAAAGCACCTTGAAGGACCACTGC	60	
QY	61	TCGGAGGACACACCAAGCGTCTGCAATGCTTACTACAGGAGGTGGACTACTGCTC	120	
Db	61	TCGGAGGACACACCAAGCGTCTGCAATGCTTACTACAGGAGGTGGACTACTGCTC	120	
QY	121	GGAGGAGGTGAGTGGCGCCGCGCTTCGGCCGCGGCGGCGGCGGCGGCGGCGGCGG	180	
Db	121	GGAGGAGGTGAGTGGCGCCGCGCTTCGGCCGCGGCGGCGGCGGCGGCGGCGGCGG	180	
QY	181	GCACGTCGTCAGGAGAAAGTTTCGAGGAGGTGCACACGCTATCACGCGCGCGGCGGCA	240	


```

RESULT 5
US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990

```

```

Query Match      25.0%; Score 169; DB 9; Length 524;
Best Local Similarity 100.0%; Pred. No. 3.5e-73;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      237  ACCACCAACCACCAACCATGGTACACACGGCGGCCACGGCTTCGTGGTGGCGGAGACCAAGGG 296
      |||||

```

Db 208 ACCACCACCACCATGTCACCGCGCGGCGCCACCGCTTCGTGGTGGAGACGAGG 267
QY 297 TCGAGGAGACATCAACACCTGCACCGCGGAGGTCCACGAGCGACGAGAGCTTCCTCG 356
Db 268 TCGAGGAGACATCAACACCTGCACCGCGGAGGTCCACGAGCGACGAGAGCTTCCTCG 327
QY 357 CAGGCGCTAATGAGCGCGCGCGCGCGGATCCACGCGCGTTCGTGC 405
Db 328 CCAGGCGTAATGAGCGCGCGCGCGCGCATCCACGCGCGTTCGTGC 376

RESULT 7

US-10-125-540-76
; Sequence 76, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-540-76

Query Match 17.9%; Score 121; DB 9; Length 2645;
Best Local Similarity 100.0%; Pred. No. 9.6e-50;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ATCAAA 615
Db 2522 ATCAAA 2581
QY 616 AA 675
Db 2582 AA 2641
QY 676 A 676
Db 2642 A 2642

RESULT 8

US-09-764-870-76
; Sequence 76, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-76

Query Match 17.9%; Score 121; DB 10; Length 2645;
Best Local Similarity 100.0%; Pred. No. 9.6e-50;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 ATCAAA 615
Db 2522 ATCAAA 2581

QY 616 AA 675
Db 2582 AA 2641
QY 676 A 676
Db 2642 A 2642

RESULT 9

US-09-832-129-21
; Sequence 21, Application US/09832129
; Publication No. US20030027297A1
; GENERAL INFORMATION:
; APPLICANT: Fiscella et al.
; TITLE OF INVENTION: 19 Human secreted proteins
; FILE REFERENCE: P2045P1
; CURRENT APPLICATION NUMBER: US/09/832,129
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: PCT/US00/28664
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/163,085
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 60/172,411
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-129-21

Query Match 17.8%; Score 120; DB 9; Length 1097;
Best Local Similarity 100.0%; Pred. No. 3.2e-49;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 TCACAA 616
Db 977 TCACAA 1036
QY 617 AA 676
Db 1037 AA 1096

RESULT 10

US-09-948-820-38
; Sequence 38, Application US/09948820
; Publication No. US20030050460A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2034P1
; CURRENT APPLICATION NUMBER: US/09/948,820
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US/09/565,391
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US99/26409
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/108,207
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-948-820-38

Query Match 17.8%; Score 120; DB 9; Length 1457;
Best Local Similarity 100.0%; Pred. No. 3.2e-49;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAA 616
DB 1331 TCAAA 1390
QY 617 AAA 676
DB 1391 AAA 1450

RESULT 11
US-10-239-676-106/c
; Sequence 106, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239, 676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 106
; LENGTH: 15832
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (2778, 2845..2846, 2865, 2916, 9325, 10589, 10612, 10809, 10817)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10821, 10827, 10830..10831, 10839, 12552, 13433, 15042)
US-10-239-676-106

Query Match 17.8%; Score 120; DB 9; Length 15832;
Best Local Similarity 100.0%; Pred. No. 2.5e-49;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAA 616
DB 4732 TCAAA 4673
QY 617 AAA 676
DB 4672 AAA 4613

RESULT 12
US-09-960-352-11058/c
; Sequence 11058, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalegan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11058
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 47-LIB34-055-Q1-EI-D12
US-09-960-352-11058

Query Match 17.6%; Score 119; DB 10; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 558 CAA 617
DB 154 CAA 95
QY 618 AAA 676
DB 94 AAA 36

RESULT 13
US-09-764-872-709
; Sequence 709, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764, 872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 709
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-709

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 558 CAA 617
DB 17 CAA 76
QY 618 AAA 676
DB 77 AAA 135

RESULT 14
US-10-091-483-344/c
; Sequence 344, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT212C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-483-344

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	558	CAAA	617
DB	149	CAAA	90
QY	618	AA	676
DB	89	AA	31

RESULT 15
US-09-764-891-7491
; Sequence 7491, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7491
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7491

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	558	CAAA	617
DB	17	CAAA	76
QY	618	AA	676
DB	77	AA	135

Search completed: June 16, 2003, 10:24:24
Job time : 155.802 secs

/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;
QY 67 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTTGGACTACTGCTCGAGGA 126
Db 60 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTTGGACTACTGCTCGAGGA 119
QY 127 GGTGAGTGTGCGGCGGCGGCTTCCGCGGCACGCGGCGGCTTCCAGCAGCAGCT 186
Db 120 GGTGAGTGTGCGGCGGCGGCTTCCGCGGCACGCGGCGGCTTCCAGCAGCAGCT 179
QY 187 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 243
Db 180 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 239
QY 244 CCACCATGCTTCCACACGCGGCGGCTTCTGTTGTCGCGGAGACGAGGTCGAGGA 303
Db 240 CCACCATGCTTCCACACGCGGCGGCTTCTGTTGTCGCGGAGACGAGGTCGAGGA 299
QY 304 GGACATCAACACTTCCAGCGGAGTTCACAGCGGCGGAGGAGCTTCTCGCGCAGGGC 363
Db 300 GGACATCAACACTTCCAGCGGAGTTCACAGCGGCGGAGGAGCTTCTCGCGCAGGGC 359
QY 364 TAACAGCGCGCGGCGGCGGCTTCCAGCGGCTTCTGTTGTCGCGGAGCTTCTCGCTTATG 423
Db 360 TAACAGCGCGCGGCGGCGGCTTCCAGCGGCTTCTGTTGTCGCGGAGCTTCTCGCTTATG 419
QY 424 TATGCTGTGTTGACTTGTGTCAGGCTATCGTATGCTTGTGCTGCTGCTTATG 483
Db 420 TATGCTGTGTTGACTTGTGTCAGGCTATCGTATGCTTGTGCTGCTGCTTATG 479
QY 484 CTCAGCTCTGTAGCAATTAAGCTTGTGACCGGCTTGTGCTGCTGCTGCTTATG 543
Db 480 CTCAGCTCTGTAGCAATTAAGCTTGTGACCGGCTTGTGCTGCTGCTGCTTATG 539
QY 544 ACTAATACCTACATCAA 603
Db 540 ACTAATAA 599
QY 604 AAAAAA 610
Db 600 AAAAAA 606

RESULT 2
BO619315 648 bp mRNA linear EST 27-JUN-2002
LOCUS RNOSEQ606.SK.abl Salt stressed Zea mays roots cDNA library Zea
DEFINITION mays cDNA clone RNOSEQ606.SK.abl similar to No homology, mRNA
sequence.
BO619315
VERSION BO619315.1 GI:21621309
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H. J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology

University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
1. 648

FEATURES
source

BASE COUNT 172 a 183 g 100 t
ORIGIN
Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;
QY 67 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTTGGACTACTGCTCGAGGA 126
Db 60 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTTGGACTACTGCTCGAGGA 119
QY 127 GGTGAGTGTGCGGCGGCGGCTTCCGCGGCACGCGGCGGCTTCCAGCAGCAGCT 186
Db 120 GGTGAGTGTGCGGCGGCGGCTTCCGCGGCACGCGGCGGCTTCCAGCAGCAGCT 179
QY 187 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 243
Db 180 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 239
QY 244 CCACCATGCTTCCACACGCGGCGGCTTCTGTTGTCGCGGAGACGAGGTCGAGGA 303
Db 240 CCACCATGCTTCCACACGCGGCGGCTTCTGTTGTCGCGGAGACGAGGTCGAGGA 299
QY 304 GGACATCAACACTTCCAGCGGAGTTCACAGCGGCGGAGGAGCTTCTCGCGCAGGGC 363
Db 300 GGACATCAACACTTCCAGCGGAGTTCACAGCGGCGGAGGAGCTTCTCGCGCAGGGC 359
QY 364 TAACAGCGCGCGGCGGCGGCTTCCAGCGGCTTCTGTTGTCGCGGAGCTTCTCGCTTATG 423
Db 360 TAACAGCGCGCGGCGGCGGCTTCCAGCGGCTTCTGTTGTCGCGGAGCTTCTCGCTTATG 419
QY 424 TATGCTGTGTTGACTTGTGTCAGGCTATCGTATGCTTGTGCTGCTGCTTATG 483
Db 420 TATGCTGTGTTGACTTGTGTCAGGCTATCGTATGCTTGTGCTGCTGCTTATG 479
QY 484 CTCAGCTCTGTAGCAATTAAGCTTGTGACCGGCTTGTGCTGCTGCTGCTTATG 543
Db 480 CTCAGCTCTGTAGCAATTAAGCTTGTGACCGGCTTGTGCTGCTGCTGCTTATG 539
QY 544 ACTAATACCTACATCAA 603
Db 540 ACTAATAA 599
QY 604 AAAAAA 610
Db 600 AAAAAA 606

RESULT 3
BO619318 648 bp mRNA linear EST 27-JUN-2002
LOCUS RNOSEQ601.SK.abl Salt stressed Zea mays roots cDNA library Zea
DEFINITION mays cDNA clone RNOSEQ601.SK.abl similar to No homology, mRNA
sequence.
BO619318
VERSION BO619318.1 GI:21621312
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

THIS PAGE BLANK (USPTO)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

FEATURES

Location/Qualifiers
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQD01.SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN
Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 67 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 126
DB 60 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 119
QY 127 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 186
DB 120 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 179
QY 187 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 243
DB 180 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 239

QY 244 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 303
DB 240 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 299
QY 304 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 363
DB 300 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 359

QY 364 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 423
DB 360 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 419
QY 424 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 483
DB 420 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 479

QY 484 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 543
DB 480 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 539
QY 544 ACTAATACCTACATCAAAAAA 603
DB 540 ACTAATACCTACATCAAAAAA 599

QY 604 AAAAAA 610
DB 600 AAAAAA 606

RESULT 4
BQ619337

LOCUS

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source

Location/Qualifiers

1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQD01.SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 67 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 126
DB 60 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 119
QY 127 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 186
DB 120 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 179

QY 187 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 243
DB 180 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 239

QY 244 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 303
DB 240 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 299
QY 304 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 363
DB 300 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 359

QY 364 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 423
DB 360 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 419
QY 424 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 483
DB 420 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 479

QY 484 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 543
DB 480 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 539
QY 544 ACTAATACCTACATCAAAAAA 603
DB 540 ACTAATACCTACATCAAAAAA 603

BQ619337 648 bp mRNA linear EST 27-JUN-2002
RNOSEQD01.SK.ab1 Salt stressed Zea mays roots cDNA library Zea
mays cDNA clone RNOSEQD01.SK.ab1 similar to No homology, mRNA
sequence.

BQ619337
BQ619337.1 GI:21621331
EST.
Zea mays.
Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

Location/Qualifiers
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQD01.SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 67 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 126
DB 60 GACACACCAAGCGTGTGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA 119
QY 127 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 186
DB 120 GTGAGGTGGTGGCCCGCGCGGCTTCGGCCGCGCACGCGCGCGCTCCAGCACACGCT 179

QY 187 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 243
DB 180 CGTCAAGGAGAGTTCGAGGAGGTGCACACGCTATCACGCGCGCGCGCA---ACCACCA 239

QY 244 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 303
DB 240 CCACCAACCATGGTGCACACGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 299
QY 304 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 363
DB 300 GGACATCAACACCTGCACCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 359

QY 364 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 423
DB 360 TAACTGAGCGCGCGCGCGCGCTTCGGTGGTGGCGGAGACGAGGTGCGGA 419
QY 424 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 483
DB 420 TATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 479

QY 484 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 543
DB 480 CTCAGCTCTGTACGAATTCAGCAATAAGCTCGTACCTGTAATAAACTTCTTCGTAAT 539
QY 544 ACTAATACCTACATCAAAAAA 603
DB 540 ACTAATACCTACATCAAAAAA 603

THIS PAGE BLANK (USPTO)